

STUDIA ET EXEMPLA
LINGUISTICA ET PHILOLOGICA

STUDIA ET EXEMPLA
LINGUISTICA ET PHILOLOGICA
(SELP)

Series V: Lexica, Tom. VI
Editores: Klaus Trost et Ernst Hansack

Series I: Studia maiora
Series II: Studia minora
Series III: Fontes et exempla linguarum Slavicarum
Series IV: Grammaticae
Series V: Lexica

ALFRED HOLL
IVAN SULJIĆ

RÜCKKLÄUFIGES
WÖRTERBUCH
ZUR KROATISCHEN
VERBALMORPHOLOGIE

Aufbereitung
mit Datenanalyseverfahren der Informatik
(Data Mining)

S. RODERER VERLAG, REGENSBURG

2010

ISBN 978-3-89783-696-9

Vorwort des Hauptautors

Dieses Buch kombiniert Fragestellungen der Linguistik mit Methoden der Informatik, speziell die strukturelle Untersuchung verbalmorphologischer Systeme mit Methoden der automatischen Datenanalyse (des Data Mining). Adressaten sind daher Informatiker ebenso wie Linguisten (einschließlich Sprachlehrer) und interessierte Laien in Gestalt von vertieft Sprachstudierenden, die den Wunsch nach detaillierten Lernhilfen und nach Einsichten in den Sprachbau mitbringen. Um den Text für Wissenschaftler der jeweils anderen Disziplin lesbar zu halten, wird die Fachterminologie beider Seiten nur reduziert verwendet und ggf. in Fußnoten erläutert.

Meine Vorarbeiten zu dieser Thematik begannen mit meiner Dissertation (Holl 1988) über Verbalsysteme in der Latein-Romania, damals noch mit der manuellen Durchführung formalisierbarer Analysealgorithmen, weil Mitte der 1980er Jahre die Speicherdimensionen von PCs denkbar gering waren. Es folgten einige kleinere Beiträge zur Skandinavistik und Romanistik (Holl 2001, 2002 und 2003).

In den Jahren 2003/04 wurden aus Mitteln der bayerischen HighTech-Offensive und der Staedtler-Stiftung Nürnberg zwei Forschungsprojekte über die Verbalmorphologie des Russischen und des Deutschen unter Verwendung moderner Informationstechnik finanziert, deren Ergebnisse als *Tom. 4* in dieser Reihe publiziert wurden (Holl 2004). 2004 bis 2006 folgten zwei weitere von der Staedtler-Stiftung geförderte Forschungsprojekte. Die Untersuchung der alt- und neugriechischen Verbalmorphologie wurde als *Tom. 5* in dieser Reihe publiziert (Holl 2006). Ein umfassendes Werk über das schwedische Substantiv – erstmalig wurde hier die Nominalmorphologie einer Sprache zum Forschungsgegenstand –, das schwedische Verb und das englische Verb (Holl 2007) erschien in englischer Sprache an der Universität Växjö, Schweden, als Band 12 der *Reihe Mathematical Modeling in Physics, Engineering and Cognitive Sciences*.

Darauf aufbauend konnte am Fachbereich Informatik der Georg-Simon-Ohm-Fachhochschule Nürnberg im Jahre 2007 ein weiteres Forschungsprojekt als Diplomarbeit durchgeführt werden. In deren Rahmen wurde mit dem Kroatischen ein – neben dem Russischen in Holl 2004 – zweiter Vertreter der slawischen Sprachfamilie mit Werkzeugen moderner Informatik analysiert. In seiner Studie untersuchte Ivan Suljic das kroatische Verbalsystem auf der Basis von ca. 4.000 Verben (Teil III). Dem Buch ist eine CD-ROM beigegeben, um Wissenschaftlern das erfasste linguistische Material in vollem Umfang und in digitalisierter Form zur Verfügung zu stellen. Für diese Studie trage ich selbst als Projektleiter die Gesamtverantwortung.

Die Einleitung (Teil I) und die einzelsprach- und wortartenunabhängige Darstellung der Prinzipien zur Datenanalyse flexionsmorphologischer Systeme (Teil II) wurden – der besseren internationalen Zugänglichkeit wegen – in englischer Sprache erweitert und teilweise verändert aus Holl 2007 übernommen, da ohne sie ein umfassendes Verständnis für die einzelsprachliche Studie nicht vermittelt werden könnte. Teil II enthält zudem Beispiele aus der englischen Verbalmorphologie, um unsere Data-Mining-Methodik auch für Informatiker ohne Kroatischkenntnisse zu demonstrieren.

Der vielfache Dank der Autoren gilt Herrn Prof. Dr. Klaus Trost und Herrn Prof. Dr. Ernst Hansack, beide Universität Regensburg, für die Aufnahme des Buches als *Tom. 6* in ihre Reihe *STUDIA ET EXEMPLA LINGUISTICA ET PHILOLOGICA Series V: Lexica* sowie für zahlreiche Anregungen im Detail. Herrn Muhidin Kešan, Lektor für Bosnisch, Kroatisch und Serbisch an der Universität Regensburg, danken wir für die Durchsicht und weitere Hinweise.

Nürnberg, im April 2010

Alfred Holl

Inhaltsverzeichnis

VORWORT DES HAUPTAUTORS	I.5
INHALTSVERZEICHNIS	I.7
TEIL I. EINFÜHRUNG - INTRODUCTION.....	I.11
1. MOTIVATION	I.13
2. METHOD..	I.14
2.1 BASICS OF THE DATA ANALYSIS	I.14
2.2 PRE-PROCESSING, PROCESSING AND POST-PROCESSING OF THE DATA ANALYSIS	I.18
3. STRUCTURE.....	I.19
TEIL II. EINZELSPRACHUNABHÄNGIGE ANTEILE DES DATA MINING FLEXIONSMORPHOLOGISCHER SYSTEME - PARTS OF DATA MINING INDEPENDENT OF INDIVIDUAL LANGUAGES.....	II.1
1. PRE-PROCESSING OF THE DATA ANALYSIS	II.3
1.1 REPRESENTATION OF LINGUISTIC OBJECTS.....	II.4
1.2 MODIFIED ORTHOGRAPHIC CONVENTIONS.....	II.5
1.2.1 <i>Treatment of the beginning of lexical bases</i>	<i>II.5</i>
1.2.2 <i>Prefix treatment</i>	<i>II.5</i>
1.3 GATHERING OF LINGUISTIC DATA	II.12
1.3.1 <i>Gathering of lexemes</i>	<i>II.12</i>
1.3.2 <i>Recording of key forms</i>	<i>II.13</i>
1.4 DELIMITATIONS.....	II.14
1.4.1 <i>Delimitations in comparison to the syntax</i>	<i>II.14</i>

1.4.2	<i>Delimitations in comparison to the lexicon</i>	II.15
1.5	KEY FEATURES AND INFLECTION TYPES	II.16
1.5.1	<i>Key features</i>	II.16
1.5.2	<i>Inflection types</i>	II.17
1.6	DERIVATION RULES AND EXCEPTIONS	II.18
1.6.1	<i>Derivation rules</i>	II.19
1.6.2	<i>Exceptions</i>	II.19
1.7	LEXEMES WITH TWO INFLECTION TYPES	II.20
1.7.1	<i>Interpretation as two lexemes</i>	II.20
1.7.2	<i>Interpretation as one lexeme</i>	II.22
2.	PROCESSING OF THE DATA ANALYSIS	II.23
2.1	TECHNICAL REQUIREMENTS.....	II.23
2.2	THE DATA MINING CONCEPT	II.23
2.2.1	<i>Preparation - sorting algorithm</i>	II.24
2.2.2	<i>Data structure and algorithm of the data mining concept</i>	II.27
3.	POST-PROCESSING OF THE DATA ANALYSIS	II.38
3.1	TYPOGRAPHIC MARKING.....	II.39
3.2	REDUCTIONS	II.39
3.2.1	<i>Formalizable reductions</i>	II.39
3.2.2	<i>Non-formalizable reductions</i>	II.45
3.3	USE OF THE RESULTING LEXEME REGISTER.....	II.45
3.3.1	<i>Assigning an arbitrary lexeme to its paradigm cluster</i>	II.45
3.3.2	<i>Gaining linguistic information from the lexeme register</i>	II.47
TEIL III. DATA MINING DES FLEXIONSMORPHOLOGISCHEN		
SYSTEMS DES KROATISCHEN VERBSIII.1		
1. VORBEREITUNG DER DATENANALYSE.....III.3		
1.1	DARSTELLUNG SPRACHLICHER EINHEITEN.....	III.3
1.2	ABWEICHENDE ORTHOGRAPHISCHE KONVENTIONEN.....	III.3
1.2.1	<i>Behandlung des Anfangs der lexikalischen Grundform</i>	III.4

1.2.2	<i>Präfix-Behandlung</i>	III.4
1.3	ZUSAMMENSTELLUNG EINES KORPUS	III.5
1.3.1	<i>Erfassung der Lexeme</i>	III.5
1.3.2	<i>Erfassung der Schlüsselformen</i>	III.6
1.4	ABGRENZUNGEN	III.6
1.4.1	<i>Abgrenzung zur Syntax</i>	III.7
1.4.2	<i>Abgrenzung zum Lexikon</i>	III.7
1.5	SCHLÜSSELANGABEN UND FLEXIONSTYPEN.....	III.8
1.5.1	<i>Schlüsselangaben</i>	III.9
1.5.2	<i>Flexionstypen</i>	III.18
1.6	ABLEITUNGSREGELN UND AUSNAHMEFORMEN.....	III.29
1.6.1	<i>Die Ableitungsregeln</i>	III.29
1.6.2	<i>Ausnahmeformen</i>	III.44
1.7	LEXEME MIT ZWEI FLEXIONSTYPEN.....	III.45
2.	DURCHFÜHRUNG DER DATENANALYSE.....	III.46
2.1	TECHNISCHE VORAUSSETZUNGEN.....	III.46
2.2	DAS DATA-MINING-KONZEPT	III.46
2.2.1	<i>Verbereitung - Sortieralgorithmus</i>	III.46
2.2.2	<i>Datenstruktur und Algorithmus des Data-Mining-Konzepts</i>	III.47
3.	NACHBEREITUNG DES DATA-MINING-PROZESSES	III.48
3.1	TYPOGRAPHISCHE KENNZEICHNUNG.....	III.48
3.2	REDUKTIONEN.....	III.48
3.2.1	<i>Formalisierbare Reduktionen</i>	III.48
3.2.2	<i>Nicht formalisierbare Reduktionen</i>	III.51
3.3	VERWENDUNG DES VERBREGISTERS	III.51
3.3.1	<i>Zuordnung eines beliebigen Verbs zu seinem Mustercluster</i>	III.51
3.3.2	<i>Gewinnung linguistischer Informationen aus dem Verbreger</i>	III.51
4.	KROATISCHES VERBREGISTER.....	III.58

TEIL IV. BIBLIOGRAPHIE.....	IV.1
1. ALLGEMEINE BIBLIOGRAPHIE.....	IV.3
2. KROATISCHE BIBLIOGRAPHIE	IV.5
ANHANG: BESCHREIBUNG DES INHALTS DER CD	IV.7

Teil I.
Einführung

Part I.
Introduction

Inhalt Teil I – Contents Part I

1. MOTIVATION	I.13
2. METHOD..	I.14
2.1 BASICS OF THE DATA ANALYSIS	I.14
2.2 PRE-PROCESSING, PROCESSING AND POST-PROCESSING OF THE DATA ANALYSIS	I.18
3. STRUCTURE.....	I.19

1. Motivation

This analysis combines data mining as a discipline of computer science and linguistic analysis of morphological systems¹.

Linguistic data analysis is nothing new. When one establishes morphological, syntactical, phonological or similar rules, data analysis is necessary to analyze linguistic material (texts, grammar books, dictionaries and especially morphological systems). Every existing list of irregular verbs, for example, is the result of a manual process of data analysis.

This approach, however, attempts to put the analysis of morphological systems on a consistent formal and automated platform using methods of computer science. For that purpose, we recur to the algorithm, independent of individual languages and parts of speech, in Holl / Behrschmidt / Kühn 2004. Already in his preparatory work, the main author of our joint project had successfully used a procedure, accessible to complete formalization, as basis of an analysis and didactic presentation of the inflectional morphology of the Latin, French, Italian, Spanish, Catalanian, Portuguese, Rumanian (Holl 1988, Holl 2003) and Swedish (Holl 2001) verbs which was formalized and implemented using IT in Holl / Behrschmidt / Kühn 2004. It was applied to the Greek verb in Holl / Pavlidis / Urban 2006.

This research project aims at the linguistic analysis and didactic presentation of the Swedish noun and verb morphologies as well as of the English verb morphology; it also demonstrates the operability of the algorithm mentioned.

This project focuses on a simple graphemic (cf. II.1.1) and synchronic view since the average learner of a language should be able to use our results without dealing with phonology and language history.

¹ In this book, morphology stands for inflectional morphology.

2. Method

2.1 Basics of the data analysis

To analyze morphological systems, data mining methods are used. “Data mining literally means mining or digging in data ... The results make patterns in data recognizable, that is why data mining is also translated as detection of data patterns” (Alpar / Niedereichholz 2000: 3, translated from German by SM). One looks for unknown, hidden connections and similarities whose knowledge promises economic or scientific benefits. In the field of data analysis, there are other expressions for the term “data mining” – according to the known instability of terminology in computer science and IT – as for example “information mining” or “knowledge discovery in databases (KDD)”, but without having the possibility of an exact definition. Today, the complete domain includes a multitude of different, partly statistical methods, which are widely used to analyze large quantities of data, as for example in marketing to find regularities in customer behavior or to detect the most promising ones among potential customers.

The methods of data analysis used in this research project belong to the group of cluster analysis algorithms. In this context, **cluster** simply means a number of data records. A data record is a tuple of data items which belong together, in this context, the key forms of a lexeme², for example the stem row of a verb (e.g. *go, went, gone*) with the reference to an inflection type (represented with a pattern lexeme), if necessary, mentioning the meaning if the inflection type depends on it (e.g. *lie – lied – lied* ‘make an untrue statement’ vs. *lie – lay – lain* ‘be in a horizontal position’). It is always the aim of a cluster analysis to find clusters with similar elements (data records), in other words “to divide data

² We use the term *lexeme* as description for a “word” as it can be found in a dictionary entry. A **lexeme** is an abstract basic unit of the lexicon which can occur in different inflectional forms. It is represented by its **lexical base** (lemma); with regard to verbs, this is the present infinitive of active voice, briefly infinitive.

in such a (semi-)automatic way into categories, classes or groups (clusters) that objects within the same cluster are as similar as possible and objects of different clusters are as dissimilar as possible” (Ester / Sander 2000: 45, translated from German by SM). Similarity has to be defined according to the specific requirements.

Up until now, detailed presentations of morphological systems are products of manual cluster analysis. In terms of data mining, clusters of morphologically analogous lexemes are established if all of the lexemes with the same irregularities are listed together with every pattern lexeme or if, in a complete list of all of the lexemes of a language-part-of-speech combination, the identifier of its pattern lexeme is mentioned with every lexeme. **Morphological analogy** means the existence of the same morphological attributes, that is, the same inflectional features, in case of verbs the same stem alternation and personal endings. But what is the profit of such clusters from a didactic point of view? They are so large and unstructured that it is impossible to memorize them.

Learners of a language use different ways in their learning processes. They want and have to minimize their learning effort and would prefer to get the irregularities right from the lexical base, regarding verbs mostly the present infinitive of active voice, just as one actually recognizes the conjugation type of regular verbs in Romance languages by looking at their infinitive endings. Once an *averbo*³ has been learned with much effort, the learner attempts to get the most of his knowledge and tries to extend it to other “similar” lexemes which he “condemns” to having the same morphological features. As *tertium comparationis* to “detect” similarities, the reverse similarity is chosen.

Reverse similarity can be seen in a common ending in the lexical base. In this sense, the term “ending” is not used as a morphological category, but simply means a sequence of letters at the end of a word. The length of an end-

³ The **averbo** of a lexeme with inflection means the ordered set of its inflectional forms.

ing is defined pragmatically depending on the particular case. **An *n*-digit ending** is defined as an ending of the length *n*, that is, the trailing *n* letters of a lexical base.

Examples for the learning strategy mentioned are the English verbs *link*, *blink* and *slink*. They have similarities in the last 4 letters, in their ending *-link*. If a learner of a language would learn the word *link* with its key forms (*linked*, *linked*, *linked*) as the first one of this group, he would want to transfer its regularities to the two other ones. This transfer is correct for *blink* (*blinked*, *blinked*), but not for *slink* (*slunk*, *slunk*). The latter has to be memorized as explicit exception of the verbs with the ending *-link*.

This strategy is correct for every basic lexeme and its prefixed lexemes (with few exceptions), for regular lexemes and a part of the irregular ones (regarding German verbs about one ninth [Holl 2002: 159]), but it is not generally correct which means a considerable source of errors. This strategy can be helpful as well as misleading. Only when it fails, the learner will intensively memorize further lexemes with their inflectional features in a second step.

In this context, two competing forms of similarity occur: the reverse similarity of the lexical base (an alphabetic property) and the morphological analogy. They are not at all identical and, within a part of speech, one cannot draw conclusions from the reverse similarity of two lexical bases onto the morphological analogy of the corresponding lexemes. The analogical reasoning (conclusion from partial to complete similarity) which underlies this strategy is a common, essential – often unconscious – principle of human learning and thinking. Therefore, it cannot simply be switched off. The learner as well as the teacher of a language has to deal with it consciously. It is best to show the learner of a language from the very beginning in which cases reverse similarity implies morphological analogy and in which it does not. This is shown in Holl 2002: 151-158 in detail.

It is, therefore, necessary to add a second one to the current form of cluster analysis of morphological systems, the latter remaining valid for reference books. This second form aims at didactically dealing with the features of ana-

logical thinking on the basis of reverse similarity. Objective is the automated explicit detection of homogenous clusters belonging to the same part of speech and having the same endings of their lexical bases. A cluster is **morphologically homogenous** (briefly homogenous) if all of its lexemes are morphologically analogous. That means that for example the verbs *cling*, *fling*, *sling*, that is, those with the ending *-ling*, form a homogenous cluster since they have the same inflectional features (*-ling*, *-lung*, *-lung*).

For this form of cluster analysis, the set of all of the lexemes of a part of speech – represented by their lexical bases – is not regarded as unsorted nor viewed in the usual alphabetic sorting from the left, but as a reversely sorted set, that is, as sorted from the right in alphabetical order. In this way, lexical bases with the same ending are placed next to each other. In cluster analysis projects known from computer science, sorted base sets do not occur so that common algorithms cannot be used in morphology.

Regarding the development of cluster analysis techniques, two basic types have to be distinguished in general: “In case of most variants, either every object is grouped as a starting cluster or all of the objects are grouped as one single cluster. After that, the starting clusters are joined or the single cluster is divided. In both cases, this is done in a way that the distances between the elements within a cluster become as small as possible” (Alpar / Niedereichholz 2000: 11, translated from German by SM). In the current case, the distance requirement is very simple; it says that clusters are detected which are as big as possible and contain morphologically analogous lexemes with the same ending. The structuring of the single all-containing starting cluster is also known as top-down cluster analysis (divisive method) and the joining of similar single-elemented starting clusters as bottom-up cluster analysis (agglomerative method). In the current project, the first variant is used since its algorithms are simpler and can be implemented more easily.

As result, one gets a stable, largely objective base on which the post-processing can be set up.

2.2 Pre-processing, processing and post-processing of the data analysis

Data selection, data clearing and transformation have to be executed before the core data analysis and require most of the time: “Although the first three steps are usually algorithmically not complex, they require, according to experts, 75-85 % of the total effort in the process of data mining” (Alpar / Niedereichholz 2000: 6 f., translated from German by SM).

An even more current paper arrives at the same conclusion: “In general, data is not available in such a form that these procedures can be applied immediately. In fact, they have to be prepared which is sixty to ninety percent of the total effort” (Kruse / Borgelt 2002: 81, translated from German by SM).

This statement is confirmed with the experiences from the preparation of linguistic data. Starting from a reverse dictionary or a lexeme list of a part of speech, the segments to be analyzed and the corresponding classificatory data are digitalized, for example read with a scanner, corrected or typed manually. The revised data is transferred into a database. The classification of inflection types used in literature is analyzed and if necessary adjusted. The changes are made in the database.

A detailed description of the pre-processing of the data analysis is presented in II.1.

The following core data analysis examines clusters of lexemes, belonging to a certain language-part-of-speech combination and having the same ending, and analyzes them regarding morphological analogy. If a cluster is morphologically analogous, the objective is reached. If not, the ending is increased by one letter to the left, the cluster is divided and one continues as above.

The analysis algorithm is described in detail in II.2.

“After using data mining techniques, the results have to be interpreted, checked and evaluated” (Kruse / Borgelt 2002: 81, translated from German by SM). Redundancy, important for learning a language, has to be maintained. To avoid misunderstandings, it is useful for the learner to explicitly know all stem

alternating lexemes and all lexemes of non-productive classes, that is, classes which new lexemes are not assigned to. Therefore, every basic lexeme of these types is mentioned explicitly, not only implicitly with a representative of a homogeneous cluster.

Possibilities of post-processing are shown in II.3.

3. Structure

The following project is divided into two further parts: the part independent of individual languages and the analysis of the Croatian verb morphology.

The structure of Part II to Part III is similar so that the reader can easily see the correspondences between the parts and therefore also between the languages. All of these three parts have the following main chapters:

- Pre-processing of the data analysis
- Processing of the data analysis
- Post-processing of the data analysis.

Teil II.

Einzel Sprachunabhängige Anteile
des Data Mining
flexionsmorphologischer Systeme

Part II.

Parts of data mining independent of
individual languages

Inhalt Teil II – Contents Part II

1. PRE-PROCESSING OF THE DATA ANALYSIS	II.3
1.1 REPRESENTATION OF LINGUISTIC OBJECTS	II.4
1.2 MODIFIED ORTHOGRAPHIC CONVENTIONS	II.5
1.3 GATHERING OF LINGUISTIC DATA	II.12
1.4 DELIMITATIONS.....	II.14
1.5 KEY FEATURES AND INFLECTION TYPES	II.16
1.6 DERIVATION RULES AND EXCEPTIONS	II.18
1.7 LEXEMES WITH TWO INFLECTION TYPES	II.20
2. PROCESSING OF THE DATA ANALYSIS.....	II.23
2.1 TECHNICAL REQUIREMENTS.....	II.23
2.2 THE DATA MINING CONCEPT	II.23
3. POST-PROCESSING OF THE DATA ANALYSIS.....	II.32
3.1 TYPOGRAPHIC MARKING	II.33
3.2 REDUCTIONS	II.33
3.3 USE OF THE RESULTING LEXEME REGISTER.....	II.39

Part II deals with basic considerations and decisions valid for all research projects of this type. Its structure and definitions are used and expanded for the parts dependent on the language-part-of-speech combinations analyzed in this book. Chapter 1 deals with the preparatory part of the data mining process. Chapter 2 shows the execution of the data mining algorithm. In Chapter 3, possibilities for the post-processing of its result are examined.

1. Pre-processing of the data analysis

As mentioned in the introduction, we use data mining methods to analyze morphological systems. These methods require a well prepared, preferably complete data source in the form of a database. The preparation of the data mining process is described in detail since such a database does not exist in most cases.

First of all, the data analysis requires principal considerations and decisions as well as definitions. In 1.1, we choose between graphemic and phonemic representation. 1.2 deals with extensions of the standard orthography which are necessary for the data mining process. 1.3 has a look at the different sources from which a database can be established. In 1.4, we describe the range of our research compared to grammar and dictionary. 1.5 gives an overview of the selection of key features and of the encoding of inflection types. In 1.6, the definitions for describing derivation rules and the possibilities of displaying exceptions are shown. 1.7 treats the alternatives of treating lexemes with two inflection types.

We use the term *lexeme* as description for a “word” as it can be found in a dictionary entry. A **lexeme** is an abstract basic unit of the lexicon which can occur in different inflectional forms. It is represented by its **lexical base**; with regard to verbs, this is the present infinitive of active voice, briefly infinitive (cf. footnote 2 in 1.2.1).

In the part independent of individual languages, we present **examples from the English verb morphology** in order to ensure a broader understanding of

our discussions. **Our methodic procedure, however, can be applied to every combination of a written language and an inflecting part of speech.**

1.1 Representation of linguistic objects

In the beginning of a project, every linguist must define the way of representation of linguistic objects: “One of the principal decisions that must be made for every grammar is how to represent the analyzed objects, that is, the question if the verb forms should be represented phonemically or graphemically” (Kempgen 1989: 4, translated from German by SM). In his Russian research project, Kempgen decides the following: “For this grammar, the orthographic representation was chosen” (Kempgen 1989: 4, translated from German by SM).

The same decision is made by the main author in his dissertation. For our joint project, we also choose the orthographic representation as the reasons stated in Kempgen 1989 and Holl 1988 – in the same structure – have not lost their validity over the past 15 years:

1. “The greater practical relevance which is achieved with this choice” (Kempgen 1989: 4, translated from German by SM).
“Current usability and more general accessibility than phonetic-phonemic character sets after the rise of written alphabetic symbols and their increasing use” (Holl 1988: 171, translated from German by SM).
2. The extensive regularity of graphophonemic relations, i.e. the rule that “a phonemic representation of the discussed problems can easily be derived from the graphemic one – but not the other way around” (Kempgen 1989: 4, translated from German by SM).
“Good approximation of the phonemic level by the graphemic one as the phonemic representation is derivable from the graphemic one using simple graphonemic rules (whereas the reverse approach would cause significant difficulties)” (Holl 1988: 171, translated from German by SM).
3. “The – compared to a phonemic notation – reduced dependency on theoretical assumptions and decisions that always assign a phonemic study to

the framework of a certain ‘paradigm’” (Kempgen 1989: 4, translated from German by SM).

”Every morphological analysis is dependent on the preceding phonemic or graphemic analysis that can go into different depths and, therefore, can deliver different results ... A generally binding, purely orthographic character set, however, that is defined by convention and tradition and has not undergone such an analysis guarantees a certain degree of independence of other linguistic disciplines” (Holl 1988: 171, translated from German by SM).

1.2 Modified orthographic conventions

In order to perform a data analysis, extensions of the orthographic conventions have to be defined. On the one hand, they affect the marking of the beginning of lexical bases (1.2.1) and, on the other hand, the special treatment of prefixes (1.2.2).

1.2.1 Treatment of the beginning of lexical bases

In order to ensure a correct execution of the data mining algorithm (2.2.2.2), it is necessary to explicitly mark the beginning of a lexical base. For this reason, it is required that the symbol # is put in front of every lexical base and is treated like a part of it. English infinitives are represented in the following way: *#run*, *#jump*, *#walk*. The hash symbol is an end-mark for the algorithm since it processes the lexical bases reversely (i.e. from right to left) letter by letter.

1.2.2 Prefix treatment

The definition under which circumstances a lexeme is called prefixed does not affect our data analysis method on the basis of reverse sorting in any way. Since minor variances can occur in the result of the data analysis, however, three possible definitions and their effects are outlined briefly:

1. Diachronic definition:

A lexeme is called prefixed if it has its origin in a compound lexeme in the history of the examined language.

2. Formal-synchronic definition:

A lexeme (foreign lexeme) is called prefixed if it is purely formally separable in prefix and basic lexeme in the current linguistic system (or in the linguistic system of the original foreign language).

3. Semantic-synchronic definition:

A lexeme is called prefixed if an educated, linguistically untrained native speaker will spontaneously understand it as a synchronically transparent composition. That is, if it is understood as a prefixed lexeme due to the similar meaning of the corresponding basic lexeme.

An example from the English language illustrates the difference between the three possible definitions:

1. Regarded from a diachronic point of view, *worship* evolved from the prefixed verb *worth-ship* by phonetic (regressive) assimilation, that is, it is prefixed according to diachronic definition.
2. Since *ship* also appears with another prefix (*re-ship*), *worship* can formal-synchronically be interpreted as prefixed as well.
3. The educated, but linguistically untrained English native speaker will spontaneously not see *worship* as a prefixed verb of *ship*, because of the significant difference in meaning between the two. Therefore, *worship* can semantic-synchronically not be considered as prefixed.

This example also demonstrates that the three definitions are compatible: prefixation according to semantic-synchronic definition implies prefixation according to formal-synchronic definition and the latter, in turn, implies prefixation according to diachronic definition.

Here is a less formal way to verbalize the set relation following from this implication: In a given language system, the set of the semantic-synchronically prefixed lexemes is a subset of the formal-synchronic ones and this one, in turn, is a subset of the diachronic ones. Or the other way round, the set of the semantic-synchronic basic lexemes is a superset of the formal-synchronic ones and this one, in turn, is a superset of the diachronic ones.

For our data analysis, we make the following decisions:

- The diachronic definition is not in our focus as we take a synchronic view from the very beginning in order to allow a learner of a language without previous knowledge of language history to understand our results.
- The decision between formal-synchronic and semantic-synchronic definition has to be made individually for every language.

An analysis on the basis of the semantic-synchronic definition of prefixation has to be judged as follows:

- For a learner of a language it is easier to work with it anyway because prefixed lexemes, which are not spontaneously transparent for educated native speakers, are even less transparent for him.
- This definition, however, is no disadvantage for the linguist either. It is true that the semantic-synchronic analysis of a lexeme inventory is obviously more difficult than the formal-synchronic one and it also creates more cases of doubt. As according to the above set relation, however, the semantic-synchronic approach will never define less basic lexemes than the formal-synchronic one. The granularity obtained will be at least as fine. This in turn means that a formal-synchronic analysis can easily be derived from a semantic-synchronic one by decomposing further lexemes. The other way round, however, is far more difficult.

Therefore, using the semantic-synchronic definition will increase the primary didactic value of our results without reducing their scientific one.

We now discuss treatment and features of prefixed lexemes.

1) Reverse sorting without prefix treatment

It is suggestive for a learner of a language to find prefixed lexemes that belong to the same basic lexeme directly after each other. This is not possible without prefix treatment as seen in the following explanations. The following form of sorting corresponds to the one used in Holl 2001. The verbs are sorted in reverse order according to Figure II.1.1. Using this type of sorting, the prefixed verbs of the basic verb *get* for instance are separated by other verbs.

. . .
get
fidget
budget
beget
parget
target
forget
. . .

Figure II.1.1: Reverse sorting without prefix treatment

In the example, *beget* and *forget*, prefixed verbs of *get*, are separated from *get* by other verbs: *beget* is between *budget* and *parget* and *forget* is listed after *target*.

2) Prefix detachment using a prefix marker

In order to avoid the described problem that occurs when sorting reversely without prefix treatment, we mark every prefix with a – (hyphen). If the hyphen is a regular grapheme in a language, one can use the _ (underscore) instead. That way, prefixed lexemes of the same basic lexeme appear immediately after each other in reversely sorted lists.

This type of presentation is significantly clearer and more helpful for the learner than the version in Figure II.1.1.

...
get
be_get
for_get
fidget
budget
parget
target
...

Figure II.1.2: Prefix marking with a hyphen

Using this procedure means a higher effort for the pre-processing part of the data mining process because the prefixed lexemes must be identified and marked with a prefix marker.

A feature is left to be discussed: There are prefixed lexemes that are not derived from a basic lexeme, at least from a synchronic point of view, in English e.g. *begin*, *recur*. There are two possibilities: to consider the prefixed lexemes as basic lexemes and not to use a hyphen or to treat them like prefixed lexemes with a basic lexeme. In this book, the second possibility is used. The following reasons support this decision:

- Prefixed lexemes with and without basic lexeme need not be distinguished any longer.
- Every language (at least every Indo-European one) has a certain collection of so-called “standard prefixes”. Regarding English verbs, these prefixes are for example *fore-*, *in-*, *inter-*, *mis-*, *over-*, *pre-*, *re-*, *un-*, *under-* etc. The knowledge of these prefixes facilitates finding the basic lexemes that namely do not exist in the language, but have to be assumed for our analysis.

Regarding this background, a treatment without the hyphen can be used for the following phenomena of morphologically inhomogeneous prefixed verbs and “pseudo-prefixed verbs” derived from nouns.

3) Morphologically inhomogeneous (pseudo-)prefixed lexemes of the same basic lexeme

Morphologically inhomogeneous (pseudo-)prefixed lexemes lead to a fragmentation of the clusters. The problem, which will be explained with regard to the English verbal system, can only be avoided using two methods:

- Pseudo-prefixed verbs¹ – derived from nouns or adjectives – that do not conjugate like their pseudo-basic verbs do not receive any prefix marker.

Examples: The regular verbs *delay*, *relay*, *behave* are pseudo-prefixed as they are derived from the nouns *delay*, *relay*, *behavior*. The regular verb *welcome* can be considered as derived from the adjective *welcome*.

- The rare prefixed verbs that do not conjugate like their basic verbs get a special prefix marker, e.g. / (bar).

Examples: The regular verbs *be//ay*, *all//ay* are regular prefixed verbs of the irregular basic verb *lay*.

The result of the data mining algorithm applied to the basic verb *lay* and its prefixed verbs illustrates this special case (Figure II.1.3). As one can see, there are four isolated clusters: *#lay* for the irregular basic verb, *_lay* for the irregular prefixed verbs (same inflection type as *#lay*), *//lay* for the regular prefixed verbs (other inflection type than *#lay*) and *elay* for the pseudo-prefixed verbs (other inflection type than *#lay*).

Using no prefix markers or inadequate prefix marking, e.g. *be_lay* and *de_lay*, would lead to more one-lexeme clusters as *be_lay* and *de_lay* would then be sorted among the morphologically homogeneous prefixed verbs of *lay*.

¹ For the German verb cf. Holl / Behrschmidt / Kühn 2004: 85 and Holl / Pavlidis / Urban 2006: II.10.

Cluster	Infinitive	Past tense.	Past participle
...			
#lay	lay	laid	laid
_lay	re_lay	relaid	relaid
	in_lay	inlaid	inlaid
	un_lay	unlaid	unlaid
	under_lay	underlaid	underlaid
	over_lay	overlaid	overlaid
	mis_lay	mislaid	mislaid
	out_lay	inlaid	inlaid
	way_lay	waylaid	waylaid
	lay	be lay	belayed
allay		allayed	allayed
elay	delay	belayed	belayed
	relay	relayed	relayed
flay	flay	flayed	flayed
play	1play	played	played
	mis_1play	misplayed	misplayed
	dis_2play	displayed	displayed
	splay	splayed	splayed
rlay	parlay	parlayed	parlayed
slay	slay	slew	slain
...			

Figure II.1.3: Example of a clustering

4) Prefixed lexemes of basic lexemes with two inflections

As their inflection is mostly determined by different meanings of the corresponding basic lexeme, as for example *lie* in the sense of ‘make an untrue statement’ (*lied, lied*) or ‘be in a horizontal position’ (*lay, lain*), prefixed lexemes can have both as well as just one of the two inflection types of the basic lexeme (cf. 1.7). We did not find any example for a prefixed lexeme with both of them. The following examples only show one of the two inflection types: *belie* (*belied, belied*) ‘give a false impression of’ or *underlie* (*underlay, underlain*) ‘lie or be situated under’.

5) Separability of prefixes

In the case of prefixed verbs, two different types can be distinguished (cf. Lühr 2000: 178 ff):

- Prefixed verbs with separable prefixes (semi-prefixes), for example *absenden (er sandte ab)*.
This phenomenon does not exist in English, therefore we quote a German example.
- Prefixed verbs with non-separable prefixes, for example *behold (he beheld)*.

The first type occurs in the German and in Scandinavian languages. The other one is the rule for Indo-European languages.

1.3 Gathering of linguistic data

As mentioned in 1.2.2, a data source in form of a database is required for the data mining analysis. If there is not any, one must create it. For this purpose, it is necessary to gather the data, that are to be analyzed, from various sources.

1.3.1 Gathering of lexemes

This section presents a short overview over the qualitatively different data sources available.

1.3.1.1 Ideal: reverse dictionary with specification of inflection

The ideal case for the preparation of a data analysis is a reverse dictionary with information about inflection and part of speech. Based on that, it is easiest to create a complete data source. A good example of such a dictionary is Zaliznjak 1987, where the major part of the Russian vocabulary (lexeme inventory) is listed with the corresponding morphological categories.

The lexemes of the part of speech examined must be extracted from the inventory of data acquired that way. One has to pay attention to the fact that lexical bases with the same ending often do not belong to the same part of

speech, i.e. the part of speech cannot clearly be identified with a glance at the ending, cf. English nouns, such as *perspective*, *adjective*, vs. English verbs, such as *to live*, *to give*, vs. English adjectives, such as *active*, *passive*, all of which have the ending *-ive*.

1.3.1.2 Almost ideal: complete alphabetical lexeme list of a part of speech with specification of inflection

A complete alphabetical lexeme list or a dictionary of the part of speech examined with specification of inflection can also be a good source. After digitalization, it can easily be sorted in reverse order. Examples are the Bescherelle editions for the French and Spanish verbs.

1.3.1.3 Not ideal: incomplete alphabetical lexeme list of a part of speech with specification of inflection

The not ideal case is encountered if there is only an incomplete alphabetical lexeme list of the part of speech examined with specification of inflection. An example is Einberger 2000 for the English verb. It can only be used in an extended form because it does not list all of the verbs.

To complete an incomplete lexeme list, it is useful to first compare the “irregular” lexemes mentioned with a list of such and add the missing ones. Secondly, it is necessary to extend it with the help of a reverse dictionary – an example of an English one is Muthmann 1999, of a German one Mater 1965. This is necessary to avoid errors in clustering. Example: the cluster *lay* would be considered as homogenous without further consideration of *belay*, *delay*, *flay*, *play* etc. (cf. 1.2.2).

1.3.2 Recording of key forms

In this section, the possibilities of recording key forms are described (cf. 1.5.1).

1.3.2.1 Manual recording

An easy, but also exhausting way to list key forms is to scan or type them.

1.3.2.2 Generation of key forms by means of an inflection type table

A more complicated way to record key forms is to generate them by means of a table in which all of the existing inflection types (cf. 1.5.2) and their corresponding key form endings are listed (cf. 2.2.2.1). The key forms themselves are generated using IT: the ending of the lexical base is replaced by the endings of the key forms of its inflection type (cf. Holl 2004 / Behrschmidt / Kühn: 206-208 for the Russian verb).

1.4 Delimitations

In order to limit the extent of our research project, we define the range of our discussions. In 1.4.1, our data analysis is delimited in comparison to the syntax and in 1.4.2 in comparison to the lexicon.

1.4.1 Delimitations in comparison to the syntax

This is only necessary in examinations of verb morphology as verbs have synthetic (single-word) and analytic (multiple-word) forms. The latter are regular and, therefore, they are not considered in our context:

“Regarding all of the forms of these verbs, the analytic ones are not considered” (Kempgen 1989: 3, translated from German by SM). For, according to Holl 1988: 175, analytic verb forms always contain a synthetic infinite verb form. What remains, is a finite verb form of an “auxiliary verb” which in turn is recorded as a verb. If the auxiliary verb form is analytic, then it must be split up in an analogical way.

In the English language, analytic forms are composed using an auxiliary verb (*have, will, be*) and a key form (infinitive, past participle):

synthetic: *find – found – found*

analytic: *I have found, I will find, I am found*

synthetic:	<i>scream – screamed – screamed</i>
analytic:	<i>I have screamed, I will scream, I am screaming</i>
synthetic:	<i>run – ran – run</i>
analytic:	<i>I have run, I will run, I am running</i>

1.4.2 Delimitations in comparison to the lexicon

Our registers do not offer complete listings of lexemes of a certain language-part-of-speech combination. A reduction to morphologically interesting material is presented which is described in 3.2 and in the particular parts dependent on individual languages.

In some languages, different auxiliary verbs are used when creating certain analytic forms. According to Holl 1988: 175 the choice of the particular auxiliary verb depends on the verb having transitive or intransitive regimen which in turn can be affected by semantic categories (for example “verbs of motion”), cf. German *ich habe gefunden* vs. *ich bin gegangen*, French *j’ai trouvé* vs. *je suis allé*. Therefore, it is suggestive to exclude the selection of auxiliary verbs from our morphological considerations and to refer to dictionaries.

A further constraint aims at the position of the accent in the lexical base. This information is not required for the execution of the data mining algorithm in general. For some languages (e.g. Greek, Russian; cf. Holl / Pavlidis / Urban 2006, Holl / Behrschmidt / Kühn 2004), however, the position of the accent must be considered both for sorting and for the data mining algorithm. In which form this is done, has to be described for the individual languages in detail.

Defective lexemes, i.e. the ones that cannot form all of the inflectional forms, “are not treated in a specific way. Only sometimes, the key form lists show whether a verb is defective or not, and if so, in which inflectional categories. Unusual or not existing key forms are marked by () or *” (Holl 1988: 175, translated from German by SM). In individual cases, a dictionary has to be considered with regard to further details.

1.5 Key features and inflection types

In this section, the possibilities to choose key features (1.5.1) and to represent inflection types (1.5.2) are explained.

1.5.1 Key features

“To reduce the linguistic material to be listed, rules (**synthesis rules**) have to be admitted. They must make it possible to derive all of the not explicitly recorded inflectional forms from the explicitly recorded key forms.

From this point of view, a morphological model is an **algorithm** (an ordered set of **orders**) which produces inflectional forms ...

Algorithm orders can be classified. We adopt the traditional division in

- (1) **constant orders** = **data** (= key forms) and
- (2) **procedure orders** = **instructions** (= synthesis rules).

Before the background of this differentiation, morphological models can basically differ with regard to two characteristic properties:

- (1) the number of key forms used and
- (2) the number and type of synthesis rules used.

... It is immediately obvious that both of the numbers are inversely proportional, that is, the fewer key forms are admitted, the more and also the more complex synthesis rules are needed” (Holl 1988: 35, translated from German by SM).

If only the lexical base were chosen as a key form, rather complex rules (process rules) would have to be invented in order to derive all of the other (possibly vowel-changing) stems and inflectional forms from it. This means: the more key forms are chosen, the more easily the derivation rules can be designed. In practice, a key form number of less than 10 proved to be absolutely sufficient in the case of verbs (cf. Holl 1988, Holl 2001, Holl / Behrschmidt / Kühn 2004, Holl / Pavlidis / Urban 2006).

The selection of the key forms (**principal parts**) is determined by the grammatical tradition of every language and developed due to practical usability. With regard to the English and German verbs, the following selections are successful (cf. III.1.5.1):

1. English

infinitive (lexical base),
past tense and
past participle

2. German

infinitive (lexical base),
past tense and
past participle as well as

3rd singular of present tense in the case of those verbs whose infinitive stem does not coincide with the present stem.

Because of the relation to the lexicon, the lexical base should always be one of the key forms. Besides key forms, other key features, such as accentuation features which are necessary for the Ancient Greek and Russian verbs, can occur as well (cf. 1.4.2).

1.5.2 Inflection types

There are several methods to label and mark morphological classes (in our research project denoted as *inflection types* in general or, in the case of verbs, as *conjugation types* or, in the case of nouns, as *declension types* in particular). The methods are introduced in the following paragraphs and can be used in pure form as well as in combination.

Artificial encoding

One of the most frequently used marking methods is artificial encoding (cf. Weermann 2001 as an example). Every morphological class gets its own number, if necessary complemented with letters or special characters. It is explicitly stated together with each lexeme listed.

Mnemonic encoding

In the case of mnemonic encoding, a brief label is used for each morphological class. It is memorable and allows associations to the type of the individual features.

Mnemonic encoding can also be done in a semi-automatic way if we compare – for every lexeme – all of its key forms (except its lexical base) to its lexical base and hence derive their differences (mathematical “distances”, e.g. in form of derivation rules). Thus we obtain inflection types dependent on key forms. If we – for every lexeme – put its “key-form-dependent inflection types” in a row, the sequence

(distance (lexical base, key form 2), ..., distance (lexical base, key form n))

can be considered as a formal mnemonic description of the lexeme’s inflection type. In the form of derivation rules, the distances can also be used to generate key forms (cf. 1.3.2.2).

Explicit listing of key forms

A further form of marking is the explicit listing of key forms. This way of notation offers the learner of a language the greatest benefit as he is thus provided with an explicit pattern (paradigm) for the inflection of morphologically analogous lexemes of the part of speech examined.

1.6 Derivation rules and exceptions

Every language-part-of-speech combination requires a specific definition of key forms and derivation rules. There are, however, fundamental principles

that are valid independent of individual languages (1.6.1). Exceptions have to be treated separately (1.6.2).

1.6.1 Derivation rules

In the parts of this book which depend on individual languages, the derivation rules for the language-part-of-speech combinations examined are listed. They are based on key forms or previously derived forms (“Zwischenform” in Fig. III.9) and allow the formation of all of the forms of every lexeme.

The derivation rules are represented in form of formal concatenation rules. In order to show more complex rules, we use the German verb *spielen* as an example:

Infinitive stem = present infinitive \ominus (e)n	<i>spiel</i>
(If the second to last letter of the infinitive is an e, en will be removed, otherwise only n.)	
1 st singular of present indicative = infinitive stem \oplus e	(ich) <i>spiele</i>
2 nd singular of present indicative =	
3 rd singular of present indicative \ominus t \oplus st	(du) <i>spielst</i>
\oplus stands for the concatenation and \ominus for the decatenation.	

1.6.2 Exceptions

With key forms and derivation rules, most of the inflectional forms can be covered. The saying “No rule without exception”, however, applies in this case as well. As an example, the English verb *to be* has the personal forms (I) *am*, (you) *are*, (he/she/it) *is*, (we) *are*, (you) *are*, (they) *are* in the present indicative. As one can easily see, it is impossible to find simple derivation rules based on the key form *be*. Such exceptions have to be mentioned explicitly. The learner of a language has to memorize them. Possibilities to deal with exceptions are:

listing the exceptions

1. separately,
2. in the register or

3. immediately after the particular inflection type.

This decision can vary from language to language and, therefore, has to be made in particular for every examination of an individual language.

1.7 Lexemes with two inflection types

As an example, there are English verbs which can belong to different inflection types. Among others, this can be determined semantically, as in the case of the verbs *lie* and *cost* (cf. Figure II.1.4), or diachronically² / diaphasically³, as for example with the archaic forms of *beget* (*begat* vs. *begot*, *begot* vs. *begotten*) and *climb* (*clomb* vs. *climbed*), or there can be no determination at all, as in *thrive* (cf. Figure II.1.5).

The phenomenon of inflectional differentiation does not play a major part within the framework of our research issues and, therefore, it is not analyzed systematically. Basically, however, there are two possibilities of interpretation. Inflection forms belonging to two different inflection types (**inflection variants**) can be interpreted as derived from two lexemes (1.7.1) or as derived from only one lexeme (1.7.2).

1.7.1 Interpretation as two lexemes

If the inflectional difference is accompanied by a semantic difference, the interpretation as two lexemes is common. As usual in dictionaries, we distinguish two such lexemes by putting a number in front of their homographic lexical bases, such as in *1lie* and *2lie* or in *1cost* and *2cost* (cf. Figure II.1.4) in the case of basic lexemes or *re_1tread* and *re_2tread* in the case of prefixed lexemes. This method is used in Parts III and IV. *cost* can be interpreted as one lexeme as well (cf. Figure II.1.6).

² depending on history (obsolete vs. common)

³ depending on situation (formal vs. colloquial)

<i>1lie</i>	<i>lied</i>	<i>lied</i>	'make untrue statement'
<i>2lie</i>	<i>lay</i>	<i>lain</i>	'be in a horizontal position'
<i>1cost</i>	<i>costed</i>	<i>costed</i>	'estimate the price'
<i>2cost</i>	<i>cost</i>	<i>cost</i>	'require payment'

Figure II.1.4: Examples of two-lexeme interpretations of inflection variants of verbs with semantic differences

If there is no semantic difference, we can use the interpretation as two lexemes as well although it is not common. In this case, we distinguish two such lexemes by putting a Greek letter in front of their lexical bases, such as in *a+thrive* and *b+thrive* (cf. Figure II.1.5). This method is used in Part IV only.

<i>a+thrive</i>	<i>thrived</i>	<i>thrived</i>	
<i>b+thrive</i>	<i>throve</i>	<i>thriven</i>	
<i>a+strive</i>	<i>strived</i>	<i>strived</i>	
<i>b+strive</i>	<i>strove</i>	<i>striven</i>	

Figure II.1.5: Examples of two-lexeme interpretations of inflection variants of verbs without further differences

The interpretation as two lexemes leads to a treatment suitable for IT which, however, is not always satisfying from a linguistic point of view as the inflection variants are separated and put in two clusters. This is always useful if the meanings are far apart, as in the case of *1lie* vs. *2lie*. If two diachronic inflection variants without different meanings, as in the case of *thrive*, shall not be separated, the interpretation as one lexeme has to be used which is done in Part III.

1.7.2 Interpretation as one lexeme

The use of the data mining algorithm requires a data source in which only one inflection type is assigned to every lexeme. The algorithm would not terminate and, therefore, a complete clustering could not be achieved if lexemes with the same lexical base would be listed twice with different inflection types.

That is why in such cases one of the inflection types must be specified as main inflection type (TYPE1 in 2.2.2.1 Figure II.2.3) which is used by the algorithm (cf. Holl 1988: 121, 123 [axiom of synonymy elimination]) and determines which cluster a lexeme is assigned to. The other inflection types (TYPE2 in 2.2.2.1, Figure II.2.3) are not considered by the data mining algorithm itself, but they appear as comments in the presentation of the results.

As far as possible, reversely similar lexemes with two inflection types each should always be assigned to the same main inflection type. In the case of the English verbs *thrive* and *strive* for example, the “regular” conjugation type is defined as main conjugation type as shown in Figure II.1.6.

<i>cost</i>	<i>costed</i> <i>cost</i>	<i>costed</i> <i>cost</i>	<i>'estimate the price'</i> <i>'require payment'</i>
<i>thrive</i>	<i>thrived</i> <i>throve</i>	<i>thrived</i> <i>thriven</i>	
<i>strive</i>	<i>strived</i> <i>strove</i>	<i>strived</i> <i>striven</i>	

Figure II.1.6: Examples of one-lexeme interpretations of inflection variants of verbs with their main conjugation type

Which inflection type is used as main inflection type, has to be defined for every particular case in research projects dependent on individual languages.

2. Processing of the data analysis

This section deals with the analysis of the linguistic material. In 2.1, the technical requirements necessary for the data analysis are described. In 2.2, our data mining concept is introduced: the preparatory algorithm and then the data structure and the functionality of the data mining algorithm in detail.

2.1 Technical requirements

A database management system (DBMS) and a compiler of a programming language or a spread-sheet program are necessary for the execution of the analysis. Compilers of various programming languages and DBMS are available in numerous variations from different producers in different price categories. It is up to the user which combination of DBMS and compiler he wants to use or whether he decides to take a spread-sheet program.

A spread-sheet program is an application whose user interface consists of a particular spread-sheet with a certain number of lines and columns. Such tools, such as Excel, are used if repeatedly occurring calculations have to be executed quickly and reliably. Spread-sheet programs do not only allow calculation, but also graphic presentation of the results. In addition, many permit access to database functions.

Experience and knowledge of the market and the complexity of the task determine the selection of the IT tools.

2.2 The data mining concept

In this section, the data mining concept is described in detail. Like every formal software concept, it consists of a data structure and an algorithm. In 2.2.1, alphabetical sorting, its problems and their solutions are described. The data structure and the algorithm are introduced from a formal point of view in 2.2.2.

2.2.1 Preparation – sorting algorithm

The reverse sorting of lexical bases with the means of IT is fairly difficult: In order to configure the data mining algorithm to be independent of individual languages, it is necessary that the sorting corresponds to the alphabetical order of the particular language. In the German language, this difficulty is recognizable with *umlaut* and *ß*.

Required sorting	Other possible sortings	
	Excel	ASCII
sagen	sagen	sagen
wagen	sägen	wagen
sägen	wagen	legen
wägen	wägen	sägen
legen	legen	wägen

Figure II.2.1: Sorting sequence in a database

As a principle, standard sorting in IT corresponds to binary encoding. This means that the sorting value of a letter corresponds to the sorting value of the number used to encode it.

Let us have a look at the column “Other possible sortings – ASCII” in Figure II.2.1. In this example from the German verb morphology, lexical bases (here infinitives) with *ä*, which are *sägen* and *wägen*, are sorted reversely after the word with *e*, which is *legen*. This is due to the ASCII encoding where the letter *ä* is sorted at the end of the alphabet.

A database only guarantees the sorting of those lexical bases which can be represented with the standard ASCII character set. The required sorting values

of non-standard characters of individual languages need not coincide with the actual sorting values of the corresponding binary numbers.

As a consequence, special attention has to be paid to correctly sorting extended character sets in other languages than English. For this purpose, it is necessary to convert an extended character set to a standard character set. Characters of the former have to be mapped via character sets of the latter onto binary numbers in such a way that required and actual sorting values coincide.

Figure II.2.2 uses the German alphabet as an example to show a conversion table. Its letters (extended character set) are assigned to characters of the ASCII character set in such a way that the lexical bases are sorted in correct order. Therefore, the order of the table shows the sorting order.

and – belong to the standard character set and are placed before numbers and letters. Therefore they do not have to be replaced.

Characters of the extended German character set	Re-encoding of the standard character set	Characters of the extended German character set	Re-encoding of the standard character set
#	#	n	F
-	-	o	G
a	1	ö	H
ä	2	p	I
b	3	q	J
c	4	r	K
d	5	s	L
e	6	ß	M
f	7	t	N
g	8	u	O
h	9	ü	P
i	A	v	Q
j	B	w	R
k	C	x	S
l	D	y	T
m	E	z	U

Figure II.2.2: Conversion table

In order to illustrate our solution, we continue with the example above: The German infinitives *wägen* and *legen* are encoded as follows:

#wägen => #R286F

#legen => #D686F

Whereas the representation with the ASCII character set does not guarantee a correct reverse sorting, the re-encoded variant is sorted correctly and, after decoding, presented in correct sequence according to the column “Required sorting” in Figure II.2.1.

In our research project, the lexemes of a language-part-of-speech combination are treated as a sorted set. They are sorted reversely according to the letter sequence of their lexical base in ascending order. This is not absolutely necessary for the data mining algorithm in its general version described in 2.2. In one run, a reference lexeme (cf. 2.2.3) is looked for until all of the lexemes with the current ending length are processed and only then the algorithm continues with the ones with the next higher ending length.

The following reasons show that it is nevertheless profitable to sort the lexemes:

1. This step is necessary for the display of the result and, therefore, has to be executed anyway.
2. A sorted set of lexemes can be searched more efficiently by a DBMS than an unsorted one.
3. The algorithm in Fig. II.2.5 can be optimized by setting a termination condition for finding numbers of comparison in the inner loop:
until `comparison_lexeme_ending (counter_lexeme_ending_length) <> current_lexeme_ending (counter_lexeme_ending_length)`.

2.2.2 Data structure and algorithm of the data mining concept

Our data mining concept consists of a data structure (2.2.2.1) and an algorithm (2.2.2.2). Both are formulated independently of parts of speech and, therefore, can be transferred directly to the special cases of the conjugation of verbs and the declension of nouns etc.

2.2.2.1 Data structure

The data structure is implemented with the help of the universal construct 'table' whether you use a DBMS or a spread-sheet program. The data structure comprises two tables.

Lexeme list

The core table “**Lexeme list**” is divided into a part independent of individual languages, necessary for the data mining algorithm and for post-processing (cf. 3), and into a part dependent on the individual language. The columns independent of individual languages are shown in Figure II.2.3.

Column name	Data type	Description
NO	num.	Unique number for each lexeme; is used as a primary key
LEXBASE	alphanum.	Lexical base
REVERSE	alphanum.	Lexical base in reverse order
SORT	alphanum.	Lexical base in reverse order and adapted encoding, cf. 2.2.1
TYPE1	alphanum.	Encoding of the main inflection type
TYPE2	alphanum.	Encoding of a further inflection type
LEXEME_ ENDING_ LENGTH	num.	Current or final lexeme ending length starting with 0
LEXEME_ CLUSTER	alphanum.	Lexeme ending ⁴ ; cluster name after the algorithm has terminated; the value NULL means that the lexeme has not been assigned to a cluster
DISPLAY	alphanum.	Specifies if a lexeme is shown in the end result of the data analysis; is used for reductions (cf. 3)

Figure II.2.3: Description of the table “Lexeme list”

⁴ We use the term *lexeme ending* as an abbreviation of the term *ending of the lexical base of a lexeme* by, so to say, identifying a lexeme with its lexical base.

The columns *LEXEME_ENDING_LENGTH* and *LEXEME_CLUSTER* are filled by the data mining algorithm. The column *DISPLAY* receives the default value *P*, that is a marker to display the data record after the analysis. During post-processing, the value can be set to *N* for non-formalizable reductions or to *F* for formalizable reductions and, therefore, on not to display the data record.

Figure II.2.3 shows only the essential elements of the data structure which are available for all language-part-of-speech combinations. Further columns necessary for a specific language-part-of-speech combination can be added any time.

Inflection types

The auxiliary table “**Inflection types**”, which can be used in addition and only serves as a comment, consists of inflection types with their encoding and a pattern lexeme each. Furthermore, a special form of this table can be used to generate key forms (cf. 1.3.2.2). The IT implementation has to be described depending on each language-part-of-speech combination examined.

Figure II.2.4 shows an example of this table for the English verb.

Column name	Data type	Description
LEXBASE	alphanum.	Lexical base of a pattern verb
TYPE	alphanum.	Encoding of the inflection type; is used as a primary key
PASTTENSE	alphanum.	Past tense 1 st singular
PASTPARTICIPLE	alphanum.	Past participle

Figure II.2.4: Complete description of the table “Inflection types”

2.2.2.2 The data mining algorithm

The flow chart in Figure II.2.5 shows the first data mining algorithm. A German example can be found in Holl / Behrschmidt / Kühn 2004: 57-73. Here we only give a brief verbal description.

The data mining algorithm leads to homogeneous clusters all of which are disjoint⁵. This means that every lexeme is assigned to exactly one cluster. Nesting and overlapping of clusters do not occur.

As already mentioned in 1.2.1, the hash sign # is used as an initial symbol of the lexical bases. This is necessary since otherwise the algorithm would not assign many lexemes to any cluster and, therefore, it would not terminate.

These facts shall be explained with the help of an example:

The English verbs *lay* and *play* form an inhomogeneous cluster with the ending length of 3 since both have the verb ending *lay*, but differ in their conjugation types. Therefore, neither is assigned to a cluster with an ending length of 3.

After increasing the ending length to 4, the verb *play* has an ending *play* whereas the ending of the verb *lay* can not be extended any further as the current ending length is larger than the verb *lay* itself. As a consequence, the verb *lay* is no longer treated by the algorithm.

In order to cope with this fact, the hash sign # is placed in front of every infinitive as its leading character to mark its beginning. Thus the verb *#lay* is treated correctly by the algorithm and is assigned to the cluster *#lay* with ending length 4.

⁵ The intersection of two clusters is always the empty set.

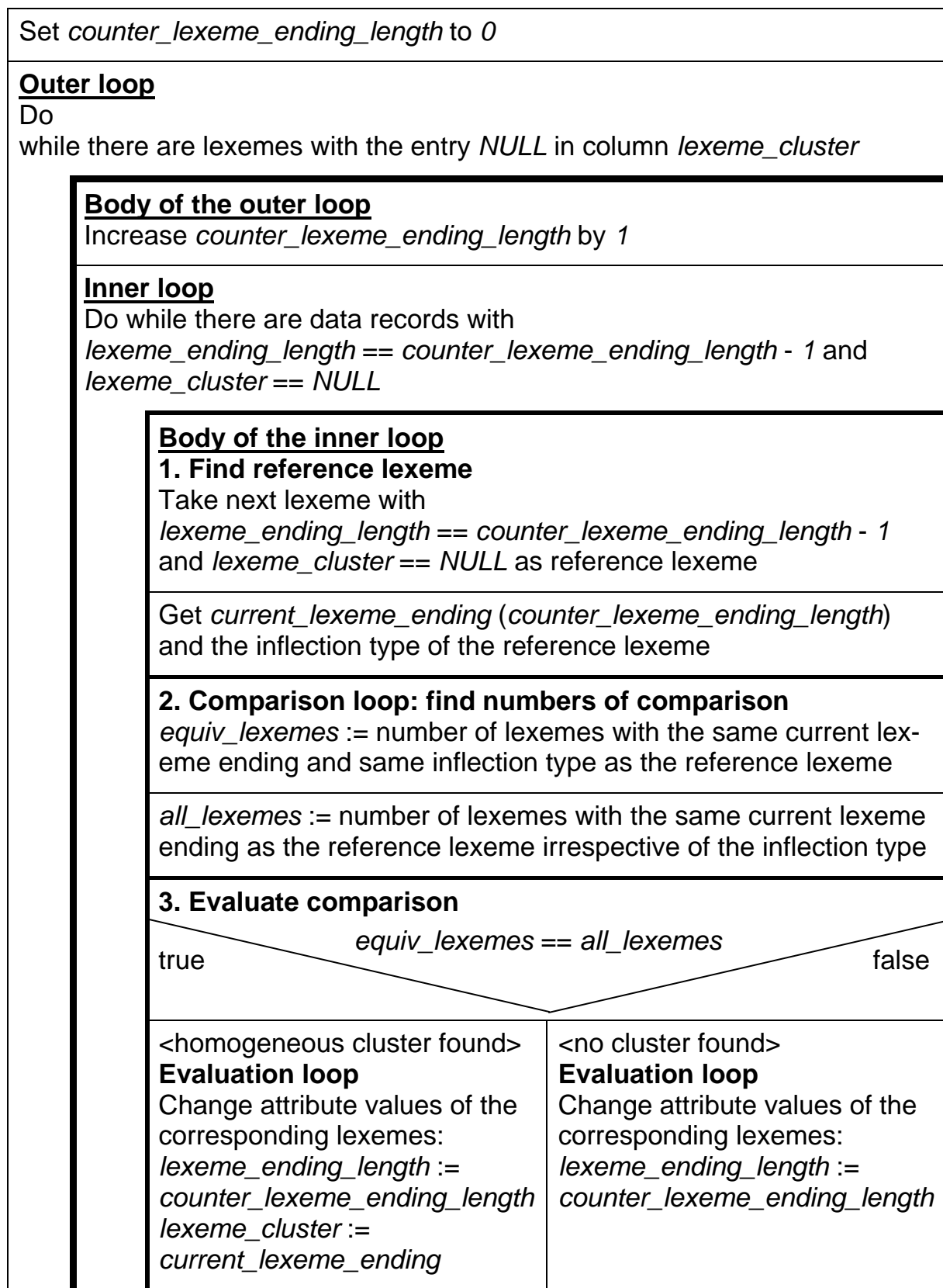


Figure II.2.5: Flow chart of the data mining algorithm I

Starting point is a lexeme list of a language-part-of-speech combination. The core of the algorithm (Fig. II.2.5) consists of two nested loops:

step by step, the outer loop processes all ending lengths;

step by step, the inner loop processes all of the lexemes, that is, all endings of a certain length.

In detail, a general step of the algorithm runs as follows.

n-th step of the outer loop:

Counter_lexeme_ending_length is increased by 1 and set to *n*.

m-th step of the inner loop:

1. The next lexeme which is not assigned to a cluster (*lexeme_ending_length* == *counter_lexeme_ending_length* - 1 and *lexeme_cluster* == NULL) is picked out. This lexeme is called **reference lexeme**. The algorithm gets the reference lexeme's ending (*current_lexeme_ending*; depending on the current value of *counter_lexeme_ending_length*) and its inflection type (**reference inflection type**).
2. In a third loop (comparison loop; not represented as a loop in the flow chart), all of the lexemes are counted which have the same *n*-digit ending, that is, the same *n* trailing letters in the lexical base, and the same inflection type as the reference lexeme. The result is calculated in the variable *equiv_lexemes*. Then, all of the lexemes are counted which have the same *n*-digit ending, no matter which inflection type they possess. The result is calculated in the variable *all_lexemes*.
3. If the variables *equiv_lexemes* and *all_lexemes* are equal, a homogeneous cluster is found. In a fourth loop (evaluation loop; not represented as a loop in the flow chart), the current lexeme ending (*current_lexeme_ending*) is stored in the column *lexeme_cluster* of the lexemes examined. If the variables are not equal, there is no homogeneous cluster. NULL remains in the column *lexeme_cluster* of the lexemes examined. In both cases, for all of

the lexemes examined, the value of the column *lexeme_ending_length* is set to *n*.

After that, the inner loop checks whether there are further reference lexemes, that is, those with *lexeme_ending_length* == *counter_lexeme_ending_length* - 1 and *lexeme_cluster* == *NULL*. In this case, the next lexeme which meets this condition becomes the next reference lexeme and the (*m*+1)-st step of the inner loop is started. Otherwise, the algorithm continues with the (*n*+1)-st step of the outer loop, that is, an increase of the current ending length by 1.

This decomposition is continued until there are only homogeneous clusters, that is, until there is no lexeme left which has the value *NULL* in the column *lexeme_cluster*.

One will often decide not to show any cluster name for single-element clusters as they do not contain any linguistic information.

The algorithm described above should use a lexeme list which is sorted reversely in alphabetic ascending order to guarantee a sorted output. Nevertheless, it does not make use of the sorting sequence and is independent of it. Therefore, the algorithm is very general and not optimized with regard to performance.

Therefore, we present a second optimized algorithm which requires a reversely sorted lexeme list (Fig. II.2.6). It makes use of the sorting sequence and runs a lot faster than the first one.

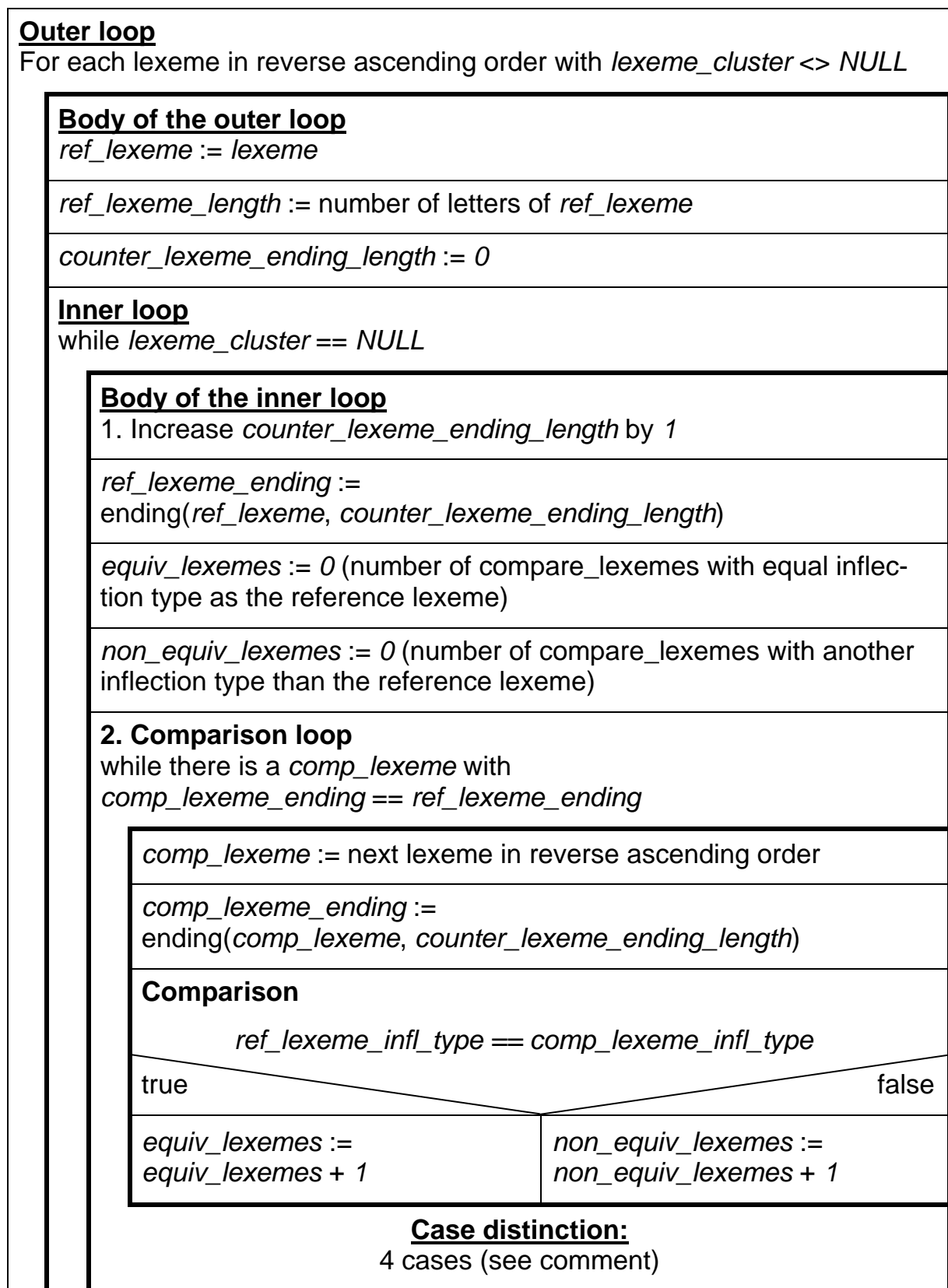


Figure II.2.6: Flow chart of the data mining algorithm II

The core of the algorithm (Fig. II.2.6) consists of two nested loops: step by step, the outer loop processes all of the lexemes; step by step, the inner loop processes all ending lengths.

In detail, a general step of the algorithm runs as follows.

n-th step of the outer loop:

The next lexeme which is not assigned to a cluster (*lexeme_cluster* == *NULL*) is picked out. This lexeme is called **reference lexeme**.

m-th step of the inner loop:

1. *Counter_lexeme_ending_length* is increased by 1 and thus set to *m*. The algorithm gets the reference lexeme's ending (*ref_lexeme_ending*, depending on the current value of *counter_lexeme_ending_length*) and its inflection type (**reference inflection type**). The algorithm is now going to compare the next lexemes with the reference lexeme. Therefore, the variables of the following comparison loop are initialized with 0.
2. In a third loop (comparison loop), all of the lexemes (*comp_lexeme*) are counted which have the same *n*-digit ending, that is, the same *n* trailing letters in the lexical base, and the same inflection type as the reference lexeme. The result is calculated in the variable *equiv_lexemes*. Then, all of the lexemes are counted which have the same *n*-digit ending and another inflection type than the reference lexeme. The result is calculated in the variable *non_equiv_lexemes*.
3. Depending on the variables *equiv_lexemes*, *non_equiv_lexemes* and *counter_lexeme_ending_length*, four cases are distinguished (Fig. II.2.7).

	<i>equiv_lexemes</i>	<i>non_equiv_lexemes</i>	<i>counter_lexeme_ending_length</i>
Case 1	0	0	—
Case 2	—	≥ 1	$\geq \text{ref_lexeme_length}$
Case 3	≥ 1	0	—
Case 4	—	≥ 1	$< \text{ref_lexeme_length}$

Figure II.2.7: Case distinction

Case 1 (there are no compare lexemes):

single-element cluster consisting of the reference lexeme only:

if *non_equivalent_lexemes* == 0 and *equivalent_lexemes* == 0
 then *lexeme_cluster* := *ref_lexeme_ending*

Case 2 (there are 1 or more compare lexemes with another inflection type than the reference lexeme and the current ending length greater or equal than the number of letters of the reference lexeme):

single-element cluster consisting of the reference lexeme only:

if *non_equivalent_lexemes* ≥ 1
 and *counter_lexeme_ending_length* $\geq \text{ref_lexeme_length}$
 then *lexeme_cluster* := *ref_lexeme*

Case 3 (all of the compare lexemes have the same inflection type as the reference lexeme):

cluster with more than 1 element:

if *non_equivalent_lexemes* == 0 and *equivalent_lexemes* > 0
 then
 for the reference lexeme
 lexeme_cluster := *ref_lexeme_ending*
evaluation loop
 for each processed *comp_lexeme*
 lexeme_cluster := *ref_lexeme_ending*

Remark: After cases 1 to 3, the inner loop terminates.

Case 4 (there are 1 or more compare lexemes with another inflection type than the reference lexeme and the current ending length less than the number of letters of the reference lexeme):

no cluster:

if *non_equivalent_lexemes* ≥ 1

and *counter_lexeme_ending_length* $<$ *ref_lexeme_length*

then no action

Remark: Only in case 4, the inner loop continues.

One will often decide not to show any cluster name for single-element clusters as they do not contain any linguistic information.

Only in case 4, that is, if no cluster was found, the inner loop continues. Its $(m+1)$ -st step starts with an increase of the current ending length (*counter_lexeme_ending_length*) by 1 and deals with a further examination of the current reference lexeme. Otherwise, that is, if a homogeneous cluster was found, the algorithm continues with the $(n+1)$ -st step of the outer loop. The next reference lexeme is examined.

This decomposition is continued until there are only homogeneous clusters, that is, until there is no lexeme left which has the value *NULL* in the column *lexeme_cluster*.

One will often decide not to show any cluster name for single-element clusters as they do not contain any linguistic information.

3. Post-processing of the data analysis

Post-processing is necessary for better evaluating the result of the data mining algorithm. Parts of the resulting list with little morphological information value have to be reduced. This reduction can have different aims, as for example absence of redundancy with maximum compression or usability in advanced language training (cf. Holl 2003: 118). The reduction, however, has to be done separately for every analyzed language-part-of-speech combination because many parts of it depend on the individual language.

Regarding the given language-part-of-speech combination, one can distinguish between formalizable and non-formalizable methods (Figure II.3.1): Formalizable ones are free of subjective influences and can be represented in form of an algorithm, non-formalizable ones are not completely objective and can be influenced by the knowledge and experience of the linguist. A part of the formalizable and all of the non-formalizable methods are dependent on individual languages as they can only be defined on the basis of individual language data.

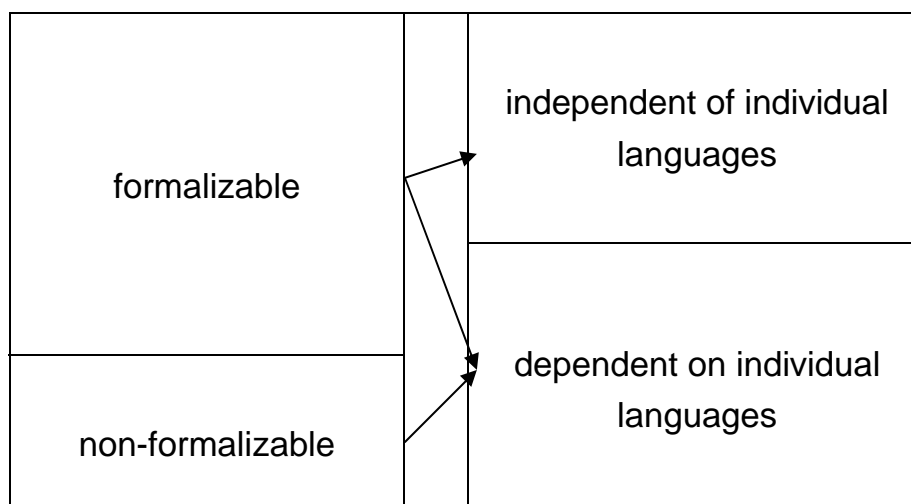


Figure II.3.1: Confrontation of formalizability and language independence

3.1 treats the typographic marking of key forms and 3.2 their reductions. 3.3 provides general comments regarding structure and usage of lexeme registers which are the result of data mining processes.

3.1 Typographic marking

Typographic marking is used to accentuate certain morphological features by means of font attributes, as for example **bold**, *italics* or underlined.

3.2 Reductions

After the execution of the data mining algorithm, the number of lexemes in some clusters is so large that the list seems to be confusing for every learner of a language. This is the main reason for significantly reducing the number of lexemes of a cluster.

3.2.1 Formalizable reductions

Formalizable reductions are reductions of the number of lexemes by means of formal methods shown in the following sections.

3.2.1.1 Reduction of prefixed lexemes in homogenous clusters

The first method to post-process the results of the data mining algorithm is the elimination of prefixed lexemes from homogenous clusters. Prefixed lexemes with the same inflection as their basic lexeme contain no linguistic information and can therefore be eliminated.

3.2.1.2 Reduction in homogenous clusters

The method described in 3.2.1.1 can not only be used for prefixed lexemes in homogenous clusters but also for homogenous clusters of “regularly inflecting” lexemes. This method, however, should only be used regarding lexemes without alternating stems to eliminate possible doubts of a learner of a language from the very beginning. A formalizable reduction could even completely remove clusters of regular lexemes from the lexeme register. This is possible without loss of information, as lexemes which are not listed always inflect regularly or are prefixed lexemes of basic lexemes in the register.

Homogeneous clusters which only comprise a single basic lexeme and its prefixed lexemes (or only a single prefixed lexeme) can always be given the name

of the lexical base of this (basic) lexeme. As this description does not contain any linguistic information, **the names of such one-lexeme clusters are not displayed** in the lexeme registers of this book.

3.2.1.3 Reduction in basic clusters

After the first run of the data mining algorithm, many homogeneous clusters show the same morphological features although they are not directly next to each other and interrupted by lexemes with different morphological features. These clusters belong to inflection types with many lexemes and are therefore joined to so-called basic clusters. They correspond to the concept of regular conjugation classes usually found in grammar books.

Every basic cluster is determined by two properties: the combination of an unequivocally defined lexeme ending (for the terminology cf. footnote 4 in 2.2.2.1) and an inflection type. For reasons of transparency, basic clusters should be disjoint, that is, not nested and not overlapping.

In order to understand basic clusters in more detail, we refine the verbal description above and introduce the concept of **connected subsets of an ordered basic set**. In mathematical set theory, a (path-)connected subset in general is a set where any two elements can be joined with a path which is entirely contained in the subset itself; connected ordered subsets do not have any gaps, non-connected ordered subsets consist of several components which are not neighbored. In our context, we only consider the special type of reversely ordered alphabet sets.

For the following definitions, we always assume some given language-part-of-speech combination, that is, some set of reversely ordered lexemes (set of lexemes reversely ordered regarding their lexical bases), as the **ordered basic set**.

An **alphabet set** (a subset of the ordered basic set) contains all of the lexemes of the connected ordered basic set which possess a certain **alphabetic property** (e.g. which have a certain ending). As our investigation is based upon re-

verse similarity, the alphabetic properties will always aim at the trailing letters of lexical bases.

Example: The alphabet set $\sim e$ comprises all of the lexemes which end in $-e$. The alphabet set $\sim e \setminus \sim oe$, comprises all of the lexemes ending in $-e$ without those ending in $-oe$.

As the basic set itself is ordered, each of its subsets is ordered. Therefore, an alphabet set is always ordered. It is **connected** if and only if it is not interrupted, that is, if it has the form of an **alphabet interval** (in our context briefly interval; cf. an interval of numbers).

As common trailing letters, that is, intervals, are in the focus of our project, we define **clusters** (not basic clusters!) as **connected alphabetic sets**.

A **non-connected** alphabet set is a **union of alphabet intervals**.

Any lexeme ending corresponds to an interval.

Example: The ending $-e$ corresponds to the interval $[-ae, -ze]$ which represents a connected alphabet set $\sim e$. The alphabet set $\sim e \setminus \sim oe$, is not connected as it is the union of the two separate intervals $[-ae, -ne]$ and $[-pe, -ze]$. Alphabet sets described with joker characters, such as C for consonant graphemes, are never connected, e.g. $\sim Ce$ which consists of all the lexemes ending in $-e$ preceded by a consonant grapheme.

Regarding this new background, we can refine our concept of a **homogeneous cluster** (cf. I.2.1). A homogeneous cluster is described with the combination of a certain **alphabetic property** (always a certain lexeme ending, that is, an interval) and a certain **morphological property** (an **inflection type**). Homogeneous clusters are **not interrupted** by lexemes with other inflection types Therefore:

Homogeneous clusters are always connected.

A **basic cluster** is described with the combination of a certain **alphabetic property** (the **alphabetic basic cluster property**; e.g. a lexeme ending) and a certain **morphological property** (an **inflection type**). A basic cluster consists of all of the lexemes which possess both of the two properties. It is the

union of homogeneous clusters with the same inflection type not all of which are neighbored, that is, which are **interrupted** by lexemes with other inflection types which lie between (outside) the homogeneous clusters. Therefore:

Basic clusters are always morphologically homogeneous and never connected.

In other words:

Connected morphologically homogeneous clusters are briefly called homogeneous clusters.

Non-connected morphologically homogeneous clusters are called basic clusters.

We already excluded nested and overlapping basic clusters. Nevertheless, we need a concept to describe some sort of embedding, a frequent necessity which is now illustrated with an example: Most of the English verbs ending in *-e* are regular and drop the *e* in the *ing*-form; their inflection type is *reg / -e* (cf. part III for details). Therefore, a basic cluster (*~e, reg / -e*) is useful. Most of the English verbs ending in *-oe*, however, do not drop the *e* in the *ing*-form; they are just regular and belong to the inflection type *reg*. The definition of a basic cluster (*~oe, reg*) appears as useful as well (cf. Part III for details). How can this situation be combined with the requirement of disjoint basic clusters?

We now introduce the two formal concepts to deal with this situation adequately.

The **basic cluster area** of a basic cluster consists of all of the lexemes which possess the alphabetic basic cluster property with no respect to the inflection type. That is, the alphabetic property is kept whereas the morphological property is dropped.

Basic cluster areas are never morphologically homogeneous.

Basic cluster areas can be connected or non-connected.

Examples: The basic cluster area *~e* is connected as it corresponds to one interval whereas *~e \ ~oe*, that is, the lexemes ending in *-e* without those ending in *-oe*, is not connected as it corresponds to the union of two separate in-

tervals:

[-ae, -ne] and [-pe, -ze].

Remark: The cluster area of a homogeneous cluster is always equal to the corresponding homogeneous cluster itself as homogeneous clusters are connected, that is, not interrupted by lexemes with the same alphabetic property and a different morphological property.

We can now formulate the restriction of nested and overlapping basic clusters more precisely:

Any pair of basic clusters has to be disjoint and any pair of basic cluster areas should be disjoint.

Exceptions with overlapping basic cluster areas are the Swedish nouns ending in *-ande* and *-an* (cf. Part IV for details).

The **connected closure of a basic cluster (of a basic cluster area)** is the lexeme interval defined with the leftmost and the rightmost lexeme of the basic cluster area (of the basic cluster area of the basic cluster):

Connected closures are never morphologically homogeneous and always connected.

Example: The connected closure of $\sim e \setminus \sim oe$ is $\sim e$.

ters are marked. If intended, even an entire basic cluster can be reduced to only one lexeme.

3.2.2 Non-formalizable reductions

Non-formalizable reductions can be influenced by the knowledge and the experience of the linguist (cf. II.3). In the course of this book, non-formalizable reductions are omitted to keep the described reductions transparent, but we will show some examples below.

Joining one-lexeme clusters to inhomogeneous clusters is an example of a non-formalizable reduction. In this case, all irregular basic lexemes have to be listed explicitly to guarantee the usability of the lexeme register, as described in 3.3.

Clusters containing lexemes with two inflection types cannot be reduced in a formalizable way. Otherwise lexemes of this type would be left out, since a reduction algorithm could not automatically respect them in homogeneous regular clusters. The linguist can also make the reduction in a way that one or two regular and all of the verbs with two conjugation types remain.

3.3 Use of the resulting lexeme register

The result of the data analysis is not some partial linguistic information, but a comprehensive overview covering the entire morphological system examined. It also provides a basis for further linguistic investigations.

In 3.3.1, instructions are given how to use a lexeme register. 3.3.2 shows ways how to obtain linguistic information from it.

3.3.1 Assigning an arbitrary lexeme to its paradigm cluster

As a principle, a lexeme register cannot claim completeness and has to be reduced as described in 3.2. Therefore, a certain procedure for the assignment of not explicitly listed lexemes to (basic or homogeneous) clusters is necessary.

The assignment follows this pattern:

1. Basic lexemes are assigned to the (basic or homogeneous) cluster whose name has the longest reverse matching. **The assigned lexeme has to be at least as long as the name of the cluster**, for example the regular English verb *peak* cannot be assigned to the irregular cluster *speak*.

If the cluster is a one-lexeme cluster (whose name is not displayed; cf. 3.2.1.2), no other lexeme can be assigned to it.

Example (Figure II.3.3): *fish* is assigned to the homogeneous cluster *~sh*.

Cluster	Inflection type	Infinitive	Past tense	Past participle
...				
sh	reg / +e	dish	dished	dished
	reg / +e	wish	wished	wished
...				

Figure II.3.3: Extract 1 from a reduced English verb register

2. **Prefixed lexemes have to be regarded in the form with a hyphenated prefix** and are otherwise treated as basic lexemes.

Example (Figure II.3.4): *mis_lay* has to be represented in hyphenated form and cannot be assigned to the cluster *slay*, but to the cluster *#lay*.

Cluster	Inflection type	Infinitive	Past tense	Past participle
...				
<i>#lay</i>	ay-ai-ai / D	lay	laid	laid
<i>/lay</i>	reg	allay	allayed	allayed
<i>elay</i>	reg	relay	relayed	relayed
<i>flay</i>	reg	flay	flayed	flayed
<i>play</i>	reg	play	played	played
<i>rlay</i>	reg	parlay	parlayed	parlayed
<i>slay</i>	ay-ew-ai / n	slay	slew	slain
...				

Figure II.3.4: Extract 2 from a reduced English verb register

3. **Only prefixed lexemes with the basic lexeme *xyz* (and the basic lexeme *xyz* itself) can be assigned to a cluster with the name *#xyz***, that is, a one-lexeme cluster (whose name is not displayed; cf. 3.2.1.2). This only happens if the prefixed lexemes do not appear explicitly in the register.

Example (Figure II.3.4): *in_lay* is assigned to the cluster *#lay* since it is a prefixed verb of *lay*. It has, therefore, the same inflection type as *lay*. The verbs *splay* and *display* are not prefixed verbs of *lay* and cannot be assigned to the cluster *#lay*, but to the cluster *play*.

The assignment rules 1, 2 and 3 always lead to an unequivocal result as long as clusters are consequently kept disjoint (cf. 2.2.2.2).

Gaining linguistic information from the lexeme register

Among others, the following possibilities are available to gain linguistic information from the data mining results:

1. Establishing a complete list of homogenous clusters with “irregular” lexemes (cf. Holl 2002).
2. Acquiring an overview of the correlations regarding morphological features within the entire system of a language-part-of-speech combination.

In our analysis, we only consider the first possibility.

Teil III.
Data Mining
des flexionsmorphologischen Systems
des kroatischen Verbs

Inhalt Teil III

1. VORBEREITUNG DER DATENANALYSE.....	III.3
1.1 DARSTELLUNG SPRACHLICHER EINHEITEN.....	III.3
1.2 ABWEICHENDE ORTHOGRAPHISCHE KONVENTIONEN	III.3
1.3 ZUSAMMENSTELLUNG EINES KORPUS	III.5
1.4 ABGRENZUNGEN	III.6
1.5 SCHLÜSSELANGABEN UND KONJUGATIONSTYPEN	III.8
1.6 ABLEITUNGSREGELN UND AUSNAHMEFORMEN.....	III.29
1.7 LEXEME MIT ZWEI FLEXIONSTYPEN.....	III.45
2. DURCHFÜHRUNG DER DATENANALYSE.....	III.46
2.1 TECHNISCHE VORAUSSETZUNGEN.....	III.46
2.2 DAS DATA-MINING-KONZEPT	III.46
3. NACHBEREITUNG DER DATENANALYSE.....	III.48
3.1 TYPOGRAPHISCHE KENNZEICHNUNG.....	III.48
3.2 REDUKTIONEN.....	III.48
3.3 VERWENDUNG DES VERBREGISTERS	III.51
4. KROATISCHES VERBREGISTER.....	III.58

Teil III behandelt die Analyse des kroatischen Verbbestandes. Die Struktur von Teil II sowie die dort getroffenen Vereinbarungen werden übernommen und für das Kroatische erweitert. 1 befasst sich mit der Vorbereitung der Datenanalyse. 2 zeigt ihre Durchführung. 3 beschreibt die Nachbereitung der Datenanalyse und 4 beinhaltet deren Resultat in Form eines Verbregisters.

1. Vorbereitung der Datenanalyse

In 1.1 werden Regelungen zur graphemischen Darstellung festgelegt. In 1.2 werden für den Data-Mining-Prozess notwendige Erweiterungen der Standard-Orthographie erläutert. In 1.3 wird das Zustandekommen der Datenbasis beschrieben. In 1.4 wird unsere Untersuchung gegenüber der Syntax und dem Lexikon abgegrenzt. In 1.5 werden die Schlüsselformen und die Konjugationstypen festgelegt. In 1.6 werden schließlich die Ableitungsregeln beschrieben und Ausnahmeformen genannt.

1.1 Darstellung sprachlicher Einheiten

Wie in II.1.1 beschrieben, beschränken wir uns auf die graphemische Darstellungsweise. Hierfür wird das kroatische Alphabet verwendet, dessen Kenntnis vorausgesetzt wird. Eine Besonderheit stellen die Kombinationen *dž*, *lj* und *nj* dar, die im Alphabet als eigenständige Buchstaben gelten. Dies muss vor allem bei der Sortierung berücksichtigt werden. Beispielsweise steht das Verb *lupiti* in der alphabetischen Reihenfolge vor dem Verb *ljubiti* (cf. 2.2.1).

1.2 Abweichende orthographische Konventionen

Außer den allgemeinen für diese Arbeit geltenden Festlegungen sind einige abweichende Konventionen von der Standard-Orthographie notwendig. Diese werden in den folgenden Abschnitten beschrieben.

1.2.1 Behandlung des Anfangs der lexikalischen Grundform

Die Behandlung des Anfangs der lexikalischen Grundform (Infinitiv) richtet sich nach der in II.1.2.1 festgelegten Konvention.

1.2.2 Präfix-Behandlung

Die in II.1.2.2 beschriebene Präfix-Behandlung wird auch für die Untersuchung der kroatischen Sprache verwendet. Der Fall der präfigierten Verben mit abtrennbarem Präfix (Halbpräfixe, Präfixoide) kommt hier nicht vor. Weitere linguistische Beschreibungskonzepte, wie Zirkumfix, Suffix und Interfix, werden nicht berücksichtigt.

1.2.2.1 Doppelpräfigierung

Im Kroatischen existiert, wie auch im Deutschen, die Besonderheit der Doppelpräfigierung. Unter Doppelpräfigierung ist das Voranstellen zweier Präfixe vor ein Simplex zu verstehen. So können beispielsweise rein formal dem Verb *laziti* „kriechen“ die Präfixe *pro-*, *na-* oder aber die Kombination *prona-* vorangestellt werden. Bei der Beurteilung von Doppelpräfigierungen können durchaus Zweifelsfälle entstehen, die durch den Linguisten zu lösen sind.

1.2.2.2 Verschmelzung an der Präfix-Simplex-Grenze

Eine Besonderheit im Kroatischen ist die grundsätzliche Vermeidung orthographischer Doppelkonsonanz. So verschmilzt das letzte Konsonantgraphem des Präfixes mit dem ersten des Simplex, wenn dieses identisch ist.

Beispielsweise bildet das Verb *sjeći* durch das Voranstellen des Präfixes *is-* das Wort *isjeći* (vs. *issjeći*). In 4 wird in diesem Fall der übrig gebliebene Buchstabe dem Simplex zugeordnet. Der Infinitiv⁹ des Verbs *isjeći* wird im Verbregister als *i-sjeći* repräsentiert.

Dies gilt nicht bei Verben, die zwei aufeinanderfolgende dunkle Vokale beinhalten, z.B. *po-oštriti*.

1.3 Zusammenstellung eines Korpus

In diesem Abschnitt beschreiben wir Literatur und Methoden für die Zusammenstellung der Datenbasis zur Untersuchung des kroatischen Verbbestandes.

1.3.1 Erfassung der Lexeme

Für das kroatische Verbalsystem existiert kein rückläufiges Wörterbuch mit Konjugationsangaben (cf. II.1.3.1.1), kein vollständiges alphabetisches Verbverzeichnis mit Konjugationsangaben (cf. II.1.3.1.2) und auch kein unvollständiges alphabetisches Verbverzeichnis mit Konjugationsangaben (cf. II.1.3.1.3). Somit muss auf andere linguistische Darstellungen zurückgegriffen werden.

1.3.1.1 Grundlage des Verbverzeichnisses

Die Verben werden dem Buch *Valenzwörterbuch kroatischer und serbischer Verben* (Panzer 2001) entnommen und von Hand erfasst. Dieses enthält nur die wichtigsten und gebräuchlichsten Verben, ist somit nicht vollständig und enthält keinerlei Angaben zur Konjugation.

⁹ Um die Flexionsformbestimmungen handlicher zu gestalten, wird die Angabe „Person“ stets weggelassen. Bei der Diathese wird „Passiv“ genannt, „Aktiv“ nicht. „Singular“ wird mit Sg. und „Plural“ mit Pl. abgekürzt. Statt „Infinitiv Präsens Aktiv“ sagen wir kurz „Infinitiv“ oder „Infinitiv Präsens“.

1.3.1.2 Erfassung der Konjugationsangaben

Die Konjugationsangaben der jeweiligen Verben sind dem Wörterbuch *Veliki rječnik hrvatskoga jezika* (Anić 2003) entnommen. Allerdings sind hier die Konjugationsangaben unvollständig. Nur Präsens 1. Sg. wird stets genannt. Bei jedem Verb werden zwar weitere Flexionsformen angegeben, diese stellen aber jeweils nur eine Auswahl dar und variieren von Verb zu Verb. Somit ist dieses Werk für den Aufbau einer Verbdatenbank nur beschränkt nutzbar. Da es für das Kroatische kein weiteres vollständiges Wörterbuch mit Konjugationsangaben gibt, muss die Zuordnung der Verben zu einer Konjugationsklasse im Rahmen dieser Untersuchung eigens durchgeführt werden.

1.3.1.3 Festlegung der Konjugationstypen

Für die Festlegung der einzelnen Konjugationstypen wird das *Grammatikhandbuch des Kroatischen unter Einschluss des Serbischen* (Kunzmann-Müller 2002) verwendet. Ein Nachteil dieses Buches ist, dass bestimmte Buchstabenalternationen innerhalb einiger Konjugationsklassen nicht behandelt werden, was weitere Konjugationsklassen erfordern würde. Die Definition zusätzlicher Konjugationsklassen muss also in dieser Untersuchung erfolgen, da keine weitere zufriedenstellende grammatische Darstellung zur Verfügung steht.

1.3.2 Erfassung der Schlüsselformen

Für die Erfassung der Schlüsselformen wurde die Methode der Generierung gewählt (siehe II.1.3.2.2). Sie wird in 1.5.2.1 ausführlich beschrieben.

1.4 Abgrenzungen

Um den Umfang der Arbeit einzuschränken, legen wir die Abgrenzungen unserer Betrachtungen fest. In 1.4.1 grenzen wir unsere Datenanalyse gegenüber der Syntax und in 1.4.2 gegenüber dem Lexikon ab.

1.4.1 Abgrenzung zur Syntax

In der Morphologie sind zwei Typen von Flexionsformen zu unterscheiden: synthetische (bestehend aus einem Wort) und analytische (bestehend aus mehr als einem Wort).

Wie auch in II.1.4.1 erläutert, werden für unsere Untersuchung die analytischen Flexionsformen nicht herangezogen. Im Kroatischen handelt es sich um das Perfekt, das Plusquamperfekt, die beiden Formen des Futurs und die beiden Konditionalformen.

Beispiel:

Infinitiv:	<i>čuvati</i>
Perfekt 1. Sg. maskulin:	<i>ja sam čuvao</i>
Plusquamperfekt 1. Sg. maskulin:	<i>bio sam čuvao</i>
Futur I 1. Sg.:	<i>čuvat ću</i>
Futur II 1. Sg. maskulin:	<i>budem čuvao</i>
Konditional I 1. Sg. maskulin:	<i>čuvao bih</i>
Konditional II 1. Sg. maskulin:	<i>bio bih čuvao</i>

1.4.2 Abgrenzung zum Lexikon

Über 4000 Verben sind in unsere Untersuchung einbezogen. Vollständigkeit kann nicht gewährleistet werden. Deshalb soll das in 4 aufgeführte Verbbregister¹⁰ nicht als Lexikonersatz dienen.

Darüber, ob eine Flexionsform bei einem bestimmten Verb tatsächlich vorkommt, gibt das Verbbregister keine Auskunft; dies muss in einem Wörterbuch nachgeschlagen werden.

¹⁰ Die Verbliste steht digital in der Datei **HR_Verbbregister.xls** zur Verfügung, die sich auf der beigefügten CD befindet. Ihr Inhalt wird im Anhang gesondert beschrieben.

1.4.2.1 Aspekte

Jedes kroatische Verb hat entweder eine perfektive (pf, z.B. *poznati*) oder imperfektive (ipf, z.B. *poznavati*) Aspektform. Zu den meisten Verben (bis auf Imperfektiva tantum und Perfektiva tantum) gibt es mindestens eine Form des jeweils anderen Aspekts. Solche Verbpaare werden Aspektpaare genannt (cf. Kunzmann-Müller 2002: 71). Selten hat ein bestimmtes Verb einen perfektiven und einen imperfektiven Aspekt. Unter welchem Aspekt ein solches Verb verwendet wird, wird erst aus dem Kontext ersichtlich.

„Der Inhalt des Verbalaspekts besteht, allgemein gesagt, darin, dass die durch das Verb gekennzeichnete Handlung hinsichtlich des Merkmals der Ganzheitlichkeit ... markiert wird“ (Kunzmann-Müller 2002: 70).

Die Aspekte sind zwar ein wichtiger Bestandteil der kroatischen Sprache, sie sind aber nicht Gegenstand unserer Untersuchung. Wir betrachten die Verben nur aufgrund ihres Infinitivs und lassen Aspekte und Aspektpaare vollständig außer Acht.

Eine nähere Erläuterung zur Kategorie des Aspekts im Kroatischen kann dem *Grammatikhandbuch des Kroatischen unter Einschluss des Serbischen* (Kunzmann-Müller 2002: 70-73) entnommen werden.

1.5 Schlüsselangaben und Flexionstypen

In diesem Abschnitt werden die für die kroatische Sprache verwendeten Schlüsselangaben ausgewählt (1.5.1) und die Flexionstypen mit deren Codierung aufgelistet (1.5.2). Dabei werden nur die in Abbildung III.1 vorgestellten synthetischen Verbalformen mit Ausnahme des Optativs betrachtet.

Da der Optativ hinsichtlich seiner Bildungsweise mit dem Partizip Präteritum formengleich ist (cf. Kunzmann-Müller 2002: 68), wird er nicht in unsere Untersuchung einbezogen. Er ist in Abbildung III.1 nur vollständigheitshalber aufgeführt. Genauere Angaben zu Funktion und Verwendung des Optativs können Kunzmann-Müller 2002: 67-68 entnommen werden.

Infinitiv	Präsens	Imperativ
<i>pjevati</i>	<i>pjevam etc.</i>	<i>pjevaj etc.</i>
<i>vući</i>	<i>vučem etc.</i>	<i>vuci etc.</i>
Aorist	Imperfekt	Partizip Präteritum
<i>pjevah etc.</i>	<i>pjevah etc.</i>	<i>pjevao etc.</i>
<i>vukoh etc.</i>	<i>vučah/vucijah etc.</i>	<i>vukao etc.</i>
Partizip Passiv	Adverbialpartizip Präteritum	Adverbialpartizip Präsens
<i>pjevan</i>	<i>pjevavši</i>	<i>pjevajući</i>
<i>vučen</i>	<i>vukavši</i>	<i>vukući</i>
Optativ		
<i>pjevao etc.</i>		
<i>vukao etc.</i>		

Abbildung III.1: Synthetischer Formenbestand

1.5.1 Schlüsselangaben

Für das Kroatische existiert keine einheitliche Festlegung von prägnanten Flexionsformen, von denen andere leicht abzuleiten sind. Im Englischen sind das typischerweise infinitive, past tense und past participle. Wir haben folgende Flexionsformen als Schlüsselformen gewählt, da sie eine optimale Ableitungs-

basis darstellen und somit die Ableitungsregeln in 1.6.1 für einen Sprachlernenden leicht zu lernen sind:

- <u>Infinitiv (lexikalische Grundform) (1.5.1.1.1)</u>	<i>brisati</i>
- <u>Präsens 3. Pl. (1.5.1.1.2)</u>	<i>brišu</i>
- <u>Partizip Präteritum Sg. feminin (1.5.1.1.3)</u>	<i>brisala</i>
- <u>Aorist 3. Sg. (1.5.1.1.4)</u>	<i>brisa</i>
- <u>Imperfekt 1. Sg. (1.5.1.1.5)</u>	<i>brisah</i>
- <u>Imperativ 2. Sg. (1.5.1.1.6)</u>	<i>briši</i>
- <u>Adverbialpartizip Präteritum (1.5.1.1.7)</u>	<i>brisavši</i>

1.5.1.1 Kommentare zu den Schlüsselformen

In diesem Abschnitt wird die Struktur der einzelnen Schlüsselformen detailliert beschrieben. Die Komplexität ihrer Bildungsregeln begründet die Entscheidung, sie als Schlüsselformen festzulegen. Es ist für den Sprachlernenden leichter, die Schlüsselform für jedes Verb zu lernen, als sie aus anderen Formen abzuleiten.

1.5.1.1.1 Strukturangaben zu Infinitiv und Lexemstamm

Der Infinitiv kann auf *-ti* oder *-ći* enden.

Die überwiegende Anzahl der kroatischen Verben endet auf *-ti*. Dem Infinitivausgang *-ti* kann sowohl einer der Vokalbuchstaben *a*, *e*, *i* oder *u* als auch der Konsonantbuchstabe *s* vorangehen. Wenn *-ti* ein Vokalbuchstabe vorangeht, wird der Lexemstamm gebildet, indem *-ti* weggelassen wird (cf. Kunzmann-Müller 2002: 79-80). Endet der Infinitiv auf *-sti*, sind zwei Fälle zu unterscheiden:

1. Der Lexemstamm endet auf *-t*, *-d*, *-s* oder *-z* und alterniert mit dem *s* im Infinitivausgang (cf. *met-* zu *mesti*, *krad-* zu *krasti*, *tres-* zu *tresti* und *muz-* zu *musti*).
2. Das Verb weist die erweiterte Infinitivendung *-sti* auf. Der Lexemstamm endet dann auf *p* oder *b* (cf. *crp-* zu *crpsti* und *greb-* zu *grepsti*).

Der Infinitivausgang *-ći* geht auf eine Verschmelzung der Infinitivendung *-ti* mit dem vorangestellten Konsonanten zurück (cf. Kunzmann-Müller 2002: 80). Der Lexemstamm kann hier auf *g*, *k*, *h* oder *đ* (bei einigen Komposita von *ići*) auslauten (cf. *strig-* zu *strići*, *pek-* zu *peći*, *vrh-* zu *vrći* und *naiđ-* zu *naići*).

Wie hier zu erkennen ist, kann bei der alleinigen Betrachtung des Infinitivs der Lexemstamm oft nicht erkannt werden. Hierfür muss die Form Präsens 3. Pl. herangezogen werden. Den Lexemstamm der Verben, die auf *-ći* oder *-sti* enden, erhält man, wenn Präsens 3. Pl. um die Personalendung *u* gekürzt wird.

Für eine kleine Menge von Verben existieren im Infinitiv Varianten, die neben der Endung *-ći* auch die Endungen *-gnuti* oder *-knuti* haben können (cf. *dignuti* und *dići*, *maknuti* und *maći*, *taknuti* und *taći*). Alle Verben, die auf *-ći* enden und eine solche Variante aufweisen, haben spezifische Flexionsformen. Nur im Präsens übernehmen sie die Flexionsformen der anderen Variante. Der Lexemstamm der auf *-ći* ausgehenden Variante endet auf *-g*, wenn dessen Präsens 3. Pl. auf *-gnu* beziehungsweise der Infinitiv der anderen Variante auf *-gnuti* auslautet, und auf *-k*, wenn dessen Präsens 3. Pl. auf *-knu* beziehungsweise der Infinitiv der anderen Variante auf *-knuti* endet.

Der Lexemstamm spielt als Zwischenform für die in 1.6.1.4 beschriebenen Ableitungsregeln des Aorists eine wichtige Rolle.

1.5.1.1.2 Strukturangaben zu Präsens 3. Pl.

Kunzmann-Müller teilt den kroatischen Verbbestand in **drei Konjugationsklassen** ein. Die Mehrheit der kroatischen Verben, deren Lexemstamm auf *-a* auslautet und die im Präsens 3. Pl. auf *-ju* enden, werden der **ersten Konjugationsklasse** zugeordnet (cf. *igrati* - *igraju*).

Zu dieser Gruppe gehören auch Verben, deren Lexemstamm auf *-je* auslautet und die im Präsens 3. Pl. auf *-ju* enden (cf. *razumjeti - razumiju*). Diese kleine Gruppe von Verben weist einen *i*-Einschub auf. Im Präsens 3. Pl. wird die Lexemstammendung *-je* weggelassen (cf. Kunzmann-Müller 2002: 43-45).

Dies wird in folgender Abbildung am Konjugationsmuster der ersten Konjugationsklasse verdeutlicht.

Infinitiv	<i>razumjeti</i>	<i>čitati</i>
Präs. 1. Sg.	<i>razumije-m</i>	<i>čita-m</i>
Präs. 2. Sg.	<i>razumije-š</i>	<i>čita-š</i>
Präs. 3. Sg.	<i>razumije-Ø</i>	<i>čita-Ø</i>
Präs. 1. Pl.	<i>razumije-mo</i>	<i>čita-mo</i>
Präs. 2. Pl.	<i>razumije-te</i>	<i>čita-te</i>
Präs. 3. Pl.	<i>razumi-ju</i>	<i>čita-ju</i>

Abbildung III.2: Konjugationsmuster der ersten Konjugationsklasse

Die erste Konjugationsklasse ist hinsichtlich der Übernahme von Wörtern aus fremden Sprachen und Wortneuschöpfungen produktiv.

Das Präsens 3. Pl. der Verben, die zur **zweiten Konjugationsklasse** gehören, endet auf *-e*. Der Lexemstamm dieser Verben endet auf *-i, -je, -lje, -nje, -re, -ča, -ža, -ja, -šta* oder *-žda*. Hierbei wird vor dem Anhängen der Personalendung *-e* der Vokal am Ende des Lexemstamms weggelassen (cf. *misliti - misle, trčati - trče, držati - drže, bojati - boje, vrištati - vrište* und *zviždati - zvižde*). Bei *-je, -lje* und *-nje* wird auch noch *j* entfernt (cf. *vidjeti - vide, željeti - žele* und *zelenjeti - zelene*) (cf. Kunzmann-Müller 2002: 45-46). Die folgende Abbildung zeigt das Konjugationsmuster der zweiten Konjugationsklasse.

Infinitiv	<i>trčati</i>	<i>vidjeti</i>
Präs. 1. Sg.	<i>trč-im</i>	<i>vid-im</i>
Präs. 2. Sg.	<i>trč-iš</i>	<i>vid-iš</i>
Präs. 3. Sg.	<i>trč-i</i>	<i>vid-i</i>
Präs. 1. Pl.	<i>trč-imo</i>	<i>vid-imo</i>
Präs. 2. Pl.	<i>trč-ite</i>	<i>vid-ite</i>
Präs. 3. Pl.	<i>trč-e</i>	<i>vid-e</i>

Abbildung III.3: Konjugationsmuster der zweiten Konjugationsklasse

In der **dritten Konjugationsklasse** werden elf Untergruppen zusammengefasst, von denen hier nur die wichtigsten genannt werden. Das Präsens 3. Pl. der Verben, die zur dritten Konjugationsklasse gehören, endet auf *-u* (cf. Kunzmann-Müller 2002: 46-51).

Die erste Unterklasse beinhaltet die überwiegende Mehrheit der Verben, deren Lexemstamm auf *-iva*, *-ova*, *-eva* oder *-uva* endet. Hierbei handelt es sich meistens um Imperfektivbildungen (cf. *kupiti* - *kupovati* und *objaviti* - *objavljivati*). Bei der Bildung des Präsens 3. Pl. wird *-ova*, *-eva* oder *-uva* durch *-uj* ersetzt. Das Präsens 3. Pl. dieser Verben endet somit auf *-uju*. Ohne Kenntnis des Infinitivs kann es hierbei zu Verwechslungen mit Verben der ersten Konjugationsklasse kommen, da dort das Präsens 3. Pl. auf *-ju* endet (cf. *trgovati* - *trguju*, *bičevati* - *bičuju* und *pljuvati* - *pljuju*).

Die zweite Unterklasse beinhaltet Verben, deren Lexemstamm auf *-a* endet und die bei der Bildung des Präsens eine Alternation aufweisen (cf. *kazati* - *kažu*, *slati* - *šalju* und *zobati* - *zoblju*). Hier kann es bei der Alternation von beispielsweise *b* nach *blj* oder *m* nach *mlj* auch zu der oben genannten Verwechslung mit Verben der ersten Konjugationsklasse kommen.

Die dritte Unterklasse beinhaltet Verben, deren Lexemstamm auf *-i* oder *-u* endet und die im Präsens einen *j*-Einschub haben (cf. *piti* - *piju* und *obuti* - *obuju*). Ohne Kenntnis des Infinitivs kann es hierbei zu Verwechslungen mit Verben der ersten Konjugationsklasse kommen, da dort das Präsens 3. Pl. ebenfalls auf *-ju* endet.

Die vierte Unterklasse beinhaltet Verben, deren Infinitiv auf *-nuti* endet. Im Präsens wird das *-u* des Lexemstamms weggelassen und die jeweilige Personalendung angehängt. Die Personalendung des Präsens 3. Pl. der dritten Konjugationsklasse lautet, wie immer, *-u* (cf. *krenuti* - *krenu* und *viknuti* - *viknu*).

Die fünfte Unterklasse beinhaltet Verben, deren Infinitiv auf *-sti* endet (cf. 1.5.1.1.1). Diese Verben zeigen Alternationen im Infinitiv (cf. *plesti* - *pletu*, und *gristi* - *grizu*). Die Alternation, die sich aus der Bildung des Präsens ergibt, muss vom Sprachlernenden auswendig gelernt werden, da sie aus dem Infinitiv nicht ableitbar ist. In dieser Gruppe stimmen Infinitiv- und Präsensstamm stets überein.

Die sechste Unterklasse beinhaltet Verben, deren Infinitiv auf *-ći* endet. Wie in 1.5.1.1.1 erläutert, endet der Lexemstamm dieser Verben auf *g*, *k* oder *h*. In allen anderen Formen des Präsens findet vor der jeweiligen Personalendung eine Alternation von *g* nach *ž*, *k* nach *č* und *h* nach *š* statt (cf. *strići* - *strižem* - *strigu*, *peći* - *pečem* - *peku* und *vrći* - *vršem* - *vrhu*).

Die restlichen Unterklassen beinhalten wenige Verben, die bei der Präsensbildung Einschübe aufweisen und somit als Sonderfälle betrachtet werden können (cf. *brati* - *beru*, *klati* - *kolju*, *početi* - *počnu*, *žeti* - *žanju*, *nasuti* - *naspu*).

Den **Präsensstamm** erhält man bei allen Verben der dritten Konjugationsklasse durch Weglassen der Personalendung *-u* im Präsens 3. Pl.

Die folgende Abbildung zeigt das Konjugationsmuster der dritten Konjugationsklasse.

Infinitiv	<i>kupovati</i>	<i>vikati</i>	<i>skoknuti</i>	<i>jesti</i>	<i>strići</i>
Präs. 1. Sg.	<i>kupuj-em</i>	<i>vič-em</i>	<i>skokn-em</i>	<i>jed-em</i>	<i>striž-em</i>
Präs. 2. Sg.	<i>kupuj-eš</i>	<i>vič-eš</i>	<i>skokn-eš</i>	<i>jed-eš</i>	<i>striž-eš</i>
Präs. 3. Sg.	<i>kupuj-e</i>	<i>vič-e</i>	<i>skokn-e</i>	<i>jed-e</i>	<i>striž-e</i>
Präs. 1. Pl.	<i>kupuj-emo</i>	<i>vič-emo</i>	<i>skokn-emo</i>	<i>jed-emo</i>	<i>striž-emo</i>
Präs. 2. Pl.	<i>kupuj-ete</i>	<i>vič-ete</i>	<i>skokn-ete</i>	<i>jed-ete</i>	<i>striž-ete</i>
Präs. 3. Pl.	<i>kupu-ju</i>	<i>vič-u</i>	<i>skokn-u</i>	<i>jed-u</i>	<i>strig-u</i>

Abbildung III.4: Konjugationsmuster der dritten Konjugationsklasse

1.5.1.1.3 Strukturangaben zum Partizip Präteritum Sg. feminin

Das Partizip Präteritum „kann von allen Verben gebildet werden. Es ist nach Genus und Numerus veränderlich“ (Engel; Mrazović 1986: 111).

Das Partizip Präteritum Sg. feminin wird gebildet, indem an den Lexemstamm *-la* angefügt wird (cf. Kunzmann-Müller 2002: 81-82). Allerdings gibt es hierzu viele Ausnahmen. Einige wichtige sollen hier genannt werden. Wenn der Lexemstamm des Verbs auf *-d* oder *-t* endet, wird vor dem Anhängen von *-la* der Konsonant weggelassen (cf. *plet-* - *plela*). Einige Verben, deren Infinitiv auf *rijeti* endet, bilden das Partizip Präteritum Sg. feminin durch Weglassen von *ije* und Anhängen von *-la* (cf. *poduprijeti* - *poduprla* und *umrijeti* - *umrla*).

Sehr wenige Verben haben mehr als zwei verschiedene Stämme (Infinitiv- und Präsensstamm). Zu diesen Verben zählen beispielsweise *ići* und seine Komposita. Es hat einen Infinitiv- und Präsensstamm *id-* (bzw. *iđ-* bei einigen Komposita), einen Aorist- und Imperfektstamm *iđ-* und einen Präteritalstamm *iš-* (bzw. *š-* bei einigen Komposita).

1.5.1.1.4 Strukturangaben zum Aorist 3. Sg.

Der Aorist 3. Sg. wird gebildet, indem an den Lexemstamm bei vokalischem Stammauslaut $-\emptyset$ und bei konsonantischem Stammauslaut $-e$ angefügt wird (cf. Kunzmann-Müller 2002: 55-57). Das heißt, bei vokalischem Stammauslaut ist der Aorist 3. Sg. gleich dem Lexemstamm (cf. *naigrati - naigra- - naigra* und *otrčati - otrč- - otrča*).

Bei Verben, deren Lexemstamm auf $-g$, $-k$ und $-h$ endet, findet im Aorist in der 2. und 3. Sg. eine Alternation von g nach $ž$, k nach $č$ und von h nach $š$ statt (cf. *reći - rek- - reče*, *dići - dig- - diže* und *vrći - vrh- - vrše*).

„Der Aorist wird vorwiegend von perfektiven Verben gebildet ... Der Aorist imperfektiver Verben ist sehr selten (etwa noch in Bibelübersetzungen) und klingt dann archaisch. Der Aorist perfektiver Verben ist in der Umgangssprache der Großstädte und auch in der Sprache der Medien ... kaum mehr üblich; hier ist er genauso wie das Imperfekt vom Perfekt verdrängt. In der Literatursprache, aber auch in der Umgangssprache der Provinz kommt der Aorist allerdings noch viel häufiger als das Imperfekt vor, er macht hier sogar dem Perfekt Konkurrenz“ (Engel; Mrazović 1986: 118 f.).

1.5.1.1.5 Strukturangaben zum Imperfekt 1. Sg.

Das Imperfekt wird aus dem Infinitiv- oder Präsensstamm gebildet. Es weist viele Alternationen auf (cf. Kunzmann-Müller 2002: 58-59).

Verben, deren Lexemstamm auf $-a$ auslautet, bilden das Imperfekt 1. Sg., indem $-h$ daran angehängt wird (cf. *igrati - igrāh* und *pjevati - pjevāh*).

Bei Verben, deren Lexemstamm auf $-g$, $-k$ oder $-h$ auslautet, wird zwischen dem Lexemstamm und der Endung $-h$ ein a eingefügt. Darüber hinaus findet eine Alternation des Lexemstammausgangs von $-g$ nach $-ž$, $-k$ nach $-č$ und von $-h$ nach $-š$ statt (cf. *strići - strig- - strižah*, *peći - pek- - pečah* und *vrći - vrh- - vršah*). Bei diesen Verben gibt es eine zweite Variante, bei der zwischen $-h$ und dem Lexemstammausgang ija eingeschoben wird. Dabei findet auch eine Alternation des Lexemstammausgangs von $-g$ nach $-z$, von $-k$ nach $-c$ und von $-h$

nach *-s* statt (cf. *peći - pek- - pecijah*). In 4 wird die gebräuchlichere erste Variante angegeben.

Verben, deren Lexemstamm auf *-i*, *-u* oder *-je* endet, bilden die Imperfektform vom Präsensstamm. Hierbei wird zwischen dem Präsensstamm und der Endung *h* ein *a* eingefügt. Darüber hinaus finden am Ausgang des Präsensstamms zahlreiche Alternationen statt, die an dieser Stelle nicht weiter behandelt werden (cf. *nositi - nos- - nošah* und *vidjeti - vid- - viđah*).

Das Imperfekt „kann nur von imperfektiven Verben gebildet werden ... [Es] kommt allerdings in der Umgangssprache und in der Sprache der Medien nicht mehr vor (hier ist es völlig vom Perfekt verdrängt). Auch in der Sprache der Literatur kommt es am ehesten noch vor, um einem Text archaischen Anstrich zu verleihen“ (Engel; Mrazović 1986: 116 f.). Das Imperfekt tritt noch als Restform in Redewendungen auf; häufiger findet es sich in der kroatischen Literatur des 19. Jh. (cf. Kunzmann-Müller 2002: 57-58).

1.5.1.1.6 Strukturangaben zum Imperativ Sg.

Der Imperativ wird meistens aus dem Präsens-, bei einigen Ausnahmen aus dem Lexemstamm gebildet.

Bei Verben der ersten Konjugationsklasse (cf. 1.5.1.1.2) wird der Imperativ Sg. gebildet, indem an deren Präsensstamm *-j* angehängt wird (cf. *igrati - igrāj* und *razumjeti - razumij*).

Bei Verben der zweiten Konjugationsklasse wird der Imperativ Sg. gebildet, indem an deren Präsensstamm *-i* angehängt wird (cf. *trčati - trči* und *vidjeti - vidi*).

Bei Verben der dritten Konjugationsklasse wird der Imperativ Sg. gebildet, indem an deren Präsensstamm *-i* oder $-\emptyset$ angehängt wird. Der Imperativ 2. Sg. der Verben, die der ersten und dritten Unterklasse der dritten Konjugationsklasse zugeordnet sind, wird gebildet, indem an deren Präsensstamm die Endung $-\emptyset$ angehängt wird (cf. *kupovati - kupuj- - kupuj* und *čuti - čuj- - čuj*). Der Imperativ 2. Sg. der Verben, die den restlichen Unterklassen der dritten Kon-

jugationsklasse zugeordnet sind, wird gebildet, indem an deren Präsensstamm die Endung *-i* angehängt wird (cf. *skoknuti - skokn- - skokni* und *kazati - kaž- - kaži*).

1.5.1.1.7 Strukturangaben zum Adverbialpartizip Präteritum

Das Adverbialpartizip Präteritum wird vom Lexemstamm abgeleitet und „wird im wesentlichen von perfektiven Verben gebildet“ (Engel; Mrazović 1986: 110).

Bei Verben mit vokalischem Lexemstammauslaut wird das Adverbialpartizip gebildet, indem an den Lexemstamm *-vši* angehängt wird (cf. *naigrati - naigravši, potražiti - potraživši* und *maknuti - maknuvši*).

Bei Verben mit konsonantischem Lexemstammauslaut wird dieses Adverbialpartizip gebildet, indem an den Lexemstamm *-avši* angehängt wird. Allerdings gibt es bei Verben mit konsonantischem Stammauslaut zahlreiche Ausnahmen, bei denen ein oder zwei Konsonanten am Lexemstammende weggelassen und *-vši* angehängt wird (cf. *sresti - sret- - srevši* und *sjesti - sjed- - sjevši*).

1.5.2 Flexionstypen

Laut Kunzmann-Müller gibt es im Kroatischen drei Konjugationsklassen (cf. 1.5.1.1.2). Die dritte Konjugationsklasse enthält weitere Unterklassen. Wir untersuchen jede Unterklasse der dritten Konjugationsklasse genauer und ordnen sie teilweise einer eigenen Flexionsklasse zu. Verben, die im Infinitiv auf *-sti* und *-ći* enden, nehmen im Kroatischen eine Sonderstellung ein und werden deshalb jeweils in zwei eigene Flexionsklassen ausgegliedert (Flexionsklasse **7** und **8**). Desweiteren werden Verben, die im Präsens einen *j*-Reflex aufweisen und daher mit Verben der ersten Konjugationsklasse verwechselt werden können, einer eigenen Flexionsklasse zugeordnet (Flexionsklasse **6**).

Darüber hinaus werden auch für Teile der ersten und zweiten Konjugationsklasse neue Flexionsklassen definiert. Die erste und zweite Konjugationsklasse enthalten Verben, die im Infinitiv auf *-jeti* enden. Da es sich hierbei um eine relativ große Anzahl von Verben handelt und bei einigen Flexionsformen eine Reihe von Alternationen auftreten, werden sie jeweils eigenen Flexionsklassen

zugeordnet (Flexionsklasse **4** und **5**). Diese Verben werden nicht in eine gemeinsame Flexionsklasse ausgegliedert, da sich deren Flexionsformen im Präsens und Imperfekt voneinander unterscheiden.

In Abbildung III.5 wird die Verfeinerung der drei Konjugationsklassen nach Kunzmann-Müller durch neu definierte Flexionsklassen dargestellt.

Diese Schritte waren nötig, um die grobgranulare Aufteilung nach Kunzmann-Müller übersichtlicher zu gestalten und die Bezeichnungen der verwendeten Flexionstypen in 1.5.2.1 zu vereinfachen.

		Konjugationsklassen nach Kunzmann-Müller		
		1	2	3
Neu definierte Flexionsklassen	1	<i>-ati</i>		
	2		<i>-iti, -ati, -eti</i>	
	3			<i>-ati, -uti, -eti</i>
	4	<i>-jeti</i>		
	5		<i>-jeti</i>	
	6			<i>-ati, -eti, -iti, -uti</i>
	7			<i>-sti</i>
	8			<i>-ći</i>

Abbildung III.5: Neu definierte Flexionsklassen

In Abbildung III.6 sind alle Flexionsklassen aufgelistet. Bei den einzelnen Einträgen handelt es sich um Ausgänge, die an einen formalen Stamm angehängt

werden. Hierbei werden die auftretenden Alternationen zwischen verschiedenen Schlüsselformen einer Flexionsklasse deutlich.

Flexionsklasse	Infinitiv	Präsens 3. Pl.	Präteritum Sg. feminin	Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpartizip Präteritum
1	<i>ati</i>	<i>aju</i>	<i>ala</i>	<i>a</i>	<i>ah</i>	<i>aj</i>	<i>avši</i>
2	<i>iti</i>	<i>e</i>	<i>la</i>	<i>i</i>	<i>ah</i>	<i>i</i>	<i>ivši</i>
	<i>biti</i>	<i>be</i>	<i>bila</i>	<i>bi</i>	<i>bljah</i>	<i>bi</i>	<i>bivši</i>
	<i>miti</i>	<i>me</i>	<i>mila</i>	<i>mi</i>	<i>mljah</i>	<i>mi</i>	<i>mivši</i>
	<i>piti</i>	<i>pe</i>	<i>pila</i>	<i>pi</i>	<i>pljah</i>	<i>pi</i>	<i>pivši</i>
	<i>viti</i>	<i>ve</i>	<i>vila</i>	<i>vi</i>	<i>vljah</i>	<i>vi</i>	<i>vivši</i>
	<i>citi</i>	<i>ce</i>	<i>cila</i>	<i>ci</i>	<i>čah</i>	<i>ci</i>	<i>civši</i>
	<i>kiti</i>	<i>ke</i>	<i>kila</i>	<i>ki</i>	<i>čah</i>	<i>ki</i>	<i>kivši</i>
	<i>diti</i>	<i>de</i>	<i>dila</i>	<i>di</i>	<i>đah</i>	<i>di</i>	<i>divši</i>
	<i>giti</i>	<i>ge</i>	<i>gila</i>	<i>gi</i>	<i>žah</i>	<i>gi</i>	<i>givši</i>
	<i>ziti</i>	<i>ze</i>	<i>zila</i>	<i>zi</i>	<i>žah</i>	<i>zi</i>	<i>zivši</i>
	<i>hiti</i>	<i>he</i>	<i>hila</i>	<i>hi</i>	<i>šah</i>	<i>hi</i>	<i>hivši</i>
	<i>siti</i>	<i>se</i>	<i>sila</i>	<i>si</i>	<i>šah</i>	<i>si</i>	<i>sivši</i>
	<i>liti</i>	<i>le</i>	<i>lila</i>	<i>li</i>	<i>ljah</i>	<i>li</i>	<i>livši</i>
	<i>niti</i>	<i>ne</i>	<i>nila</i>	<i>ni</i>	<i>njah</i>	<i>ni</i>	<i>nivši</i>
	<i>sliti</i>	<i>sle</i>	<i>slila</i>	<i>sli</i>	<i>šljah</i>	<i>sli</i>	<i>slivši</i>
	<i>titi</i>	<i>te</i>	<i>tila</i>	<i>ti</i>	<i>ćah</i>	<i>ti</i>	<i>tivši</i>
	<i>stiti</i>	<i>ste</i>	<i>stila</i>	<i>sti</i>	<i>ščah</i>	<i>sti</i>	<i>stivši</i>
	<i>stiti</i>	<i>ste</i>	<i>stila</i>	<i>sti</i>	<i>štah</i>	<i>sti</i>	<i>stivši</i>
	<i>zniti</i>	<i>zne</i>	<i>znila</i>	<i>zni</i>	<i>žnah</i>	<i>zni</i>	<i>znivši</i>
	<i>sniti</i>	<i>sne</i>	<i>snila</i>	<i>sni</i>	<i>šnjah</i>	<i>sni</i>	<i>snivši</i>
	<i>eti</i>	<i>e</i>	<i>ela</i>	<i>e</i>	--	<i>i</i>	<i>evši</i>
	<i>ati</i>	<i>e</i>	<i>ala</i>	<i>a</i>	<i>ah</i>	<i>i</i>	<i>avši</i>

Fle- xions- klasse	Infi- nitiv	Präsens 3. Pl.	Präteri- tum Sg. feminin	Aorist 3. Sg.	Imper- fekt 1. Sg.	Impe- rativ Sg.	Adverbial- partizip Präteritum
3	<i>ati</i>	<i>u</i>	<i>ala</i>	<i>a</i>	<i>ah</i>	<i>i</i>	<i>avši</i>
	<i>eti</i>	<i>nu</i>	<i>ela</i>	<i>e</i>	--	<i>ni</i>	<i>evši</i>
	<i>cati</i>	<i>ču</i>	<i>cala</i>	<i>ca</i>	<i>cah</i>	<i>či</i>	<i>cavši</i>
	<i>kati</i>	<i>ču</i>	<i>kala</i>	<i>ka</i>	<i>kah</i>	<i>či</i>	<i>kavši</i>
	<i>dati</i>	<i>đu</i>	<i>dala</i>	<i>da</i>	<i>dah</i>	<i>đi</i>	<i>davši</i>
	<i>gati</i>	<i>žu</i>	<i>gala</i>	<i>ga</i>	<i>gah</i>	<i>ži</i>	<i>gavši</i>
	<i>zati</i>	<i>žu</i>	<i>zala</i>	<i>za</i>	<i>zah</i>	<i>ži</i>	<i>zavši</i>
	<i>hati</i>	<i>šu</i>	<i>hala</i>	<i>ha</i>	<i>hah</i>	<i>ši</i>	<i>havši</i>
	<i>sati</i>	<i>šu</i>	<i>sala</i>	<i>sa</i>	<i>sah</i>	<i>ši</i>	<i>savši</i>
	<i>tati</i>	<i>ću</i>	<i>tala</i>	<i>ta</i>	<i>tah</i>	<i>ći</i>	<i>tavši</i>
	<i>eti</i>	<i>mu</i>	<i>ela</i>	<i>e</i>	--	<i>mi</i>	<i>evši</i>
	<i>uti</i>	<i>mu</i>	<i>ula</i>	<i>u</i>	--	<i>mi</i>	<i>uvši</i>
	<i>uti</i>	<i>pu</i>	<i>ula</i>	<i>u</i>	--	<i>pi</i>	<i>uvši</i>
	<i>uti</i>	<i>u</i>	<i>ula</i>	<i>u</i>	--	<i>i</i>	<i>uvši</i>

Flexions- klasse	Infinitiv	Präsens 3. Pl.	Präteritum Sg. feminin	Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbial- partizip Präteritum
4	<i>bjeti</i>	<i>biju</i>	<i>bjela</i>	<i>bje</i>	<i>bijah</i>	<i>bij</i>	<i>bjevši</i>
	<i>mjeti</i>	<i>miju</i>	<i>mjela</i>	<i>mje</i>	<i>mijah</i>	<i>mij</i>	<i>mjevši</i>
	<i>pjeti</i>	<i>piju</i>	<i>pjela</i>	<i>pje</i>	<i>pijah</i>	<i>pij</i>	<i>pjevši</i>
	<i>vjeti</i>	<i>viju</i>	<i>vjela</i>	<i>vje</i>	<i>vijah</i>	<i>vij</i>	<i>vjevši</i>
	<i>cjeti</i>	<i>ciju</i>	<i>cjela</i>	<i>cje</i>	<i>cijah</i>	<i>cij</i>	<i>cjevši</i>
	<i>kjeti</i>	<i>kiju</i>	<i>kjela</i>	<i>kje</i>	<i>kijah</i>	<i>kij</i>	<i>kjevši</i>
	<i>djeti</i>	<i>diju</i>	<i>djela</i>	<i>dje</i>	<i>dijah</i>	<i>dij</i>	<i>djevši</i>
	<i>gjeti</i>	<i>giju</i>	<i>gjela</i>	<i>gje</i>	<i>gijah</i>	<i>gij</i>	<i>gjevši</i>
	<i>zjeti</i>	<i>ziju</i>	<i>zjela</i>	<i>zje</i>	<i>zijah</i>	<i>zij</i>	<i>zjevši</i>
	<i>hjeti</i>	<i>hiju</i>	<i>hjela</i>	<i>hje</i>	<i>hijah</i>	<i>hij</i>	<i>hjevši</i>
	<i>sjeti</i>	<i>siju</i>	<i>sjela</i>	<i>sje</i>	<i>sijah</i>	<i>sij</i>	<i>sjevši</i>
	<i>jeti</i>	<i>iju</i>	<i>jela</i>	<i>je</i>	<i>ijah</i>	<i>ij</i>	<i>jevši</i>
	<i>ljeti</i>	<i>liju</i>	<i>ljela</i>	<i>lje</i>	<i>lijah</i>	<i>lij</i>	<i>ljevši</i>
	<i>njeti</i>	<i>niju</i>	<i>njela</i>	<i>nje</i>	<i>nijah</i>	<i>nij</i>	<i>njevši</i>
	<i>sljeti</i>	<i>sliju</i>	<i>sljela</i>	<i>slje</i>	<i>slijah</i>	<i>slij</i>	<i>sljevši</i>
	<i>tjeti</i>	<i>tiju</i>	<i>tjela</i>	<i>tje</i>	<i>tijah</i>	<i>tij</i>	<i>tjevši</i>

Fle- xions- klasse	Infi- nitiv	Präsens 3. Pl.	Präteri- tum Sg. feminin	Aorist 3. Sg.	Imper- fekt 1. Sg.	Impe- rativ Sg.	Adverbial- partizip Präteritum
5	<i>bjeti</i> <i>mjeti</i> <i>pjeti</i> <i>vjeti</i> <i>cjeti</i> <i>kjeti</i> <i>djeti</i> <i>gjeti</i> <i>zjeti</i> <i>hjeti</i> <i>sjeti</i> <i>jeti</i> <i>ljeti</i> <i>njeti</i> <i>sljeti</i> <i>tjeti</i>	<i>be</i> <i>me</i> <i>pe</i> <i>ve</i> <i>ce</i> <i>ke</i> <i>de</i> <i>ge</i> <i>ze</i> <i>he</i> <i>se</i> <i>e</i> <i>le</i> <i>ne</i> <i>sle</i> <i>te</i>	<i>bjela</i> <i>mjela</i> <i>pjela</i> <i>vjela</i> <i>cjela</i> <i>kjela</i> <i>djela</i> <i>gjela</i> <i>zjela</i> <i>hjela</i> <i>sjela</i> <i>jela</i> <i>ljela</i> <i>njela</i> <i>sljela</i> <i>tjela</i>	<i>bje</i> <i>mje</i> <i>pje</i> <i>vje</i> <i>cje</i> <i>kje</i> <i>dje</i> <i>gje</i> <i>zje</i> <i>hje</i> <i>sje</i> <i>je</i> <i>lje</i> <i>nje</i> <i>slje</i> <i>tje</i>	<i>bljah</i> <i>mljah</i> <i>pljah</i> <i>vljah</i> <i>čah</i> <i>čah</i> <i>đah</i> <i>žah</i> <i>žah</i> <i>šah</i> <i>šah</i> <i>jah</i> <i>ljah</i> <i>njah</i> <i>šljah</i> <i>ćah</i>	<i>bi</i> <i>mi</i> <i>pi</i> <i>vi</i> <i>ci</i> <i>ki</i> <i>di</i> <i>gi</i> <i>zi</i> <i>hi</i> <i>si</i> <i>i</i> <i>li</i> <i>ni</i> <i>sli</i> <i>ti</i>	<i>bjevši</i> <i>mjevši</i> <i>pjevši</i> <i>vjevši</i> <i>cjevši</i> <i>kjevši</i> <i>djevši</i> <i>gjevši</i> <i>zjevši</i> <i>hjevši</i> <i>sjevši</i> <i>jevši</i> <i>ljevši</i> <i>njevši</i> <i>sljevši</i> <i>tjevši</i>
6	<i>avati</i> <i>ivati</i> <i>ovati</i> <i>iti</i> <i>uti</i> <i>ati</i> <i>eti</i>	<i>aju</i> <i>uju</i> <i>uju</i> <i>iju</i> <i>uju</i> <i>lju</i> <i>nju</i>	<i>avala</i> <i>ivala</i> <i>ovala</i> <i>ila</i> <i>ula</i> <i>ala</i> <i>ela</i>	<i>ava</i> <i>iva</i> <i>ova</i> <i>i</i> <i>u</i> <i>a</i> <i>e</i>	<i>avah</i> <i>ivah</i> <i>ovah</i> <i>ijah</i> <i>ujah</i> <i>ah</i> <i>--</i>	<i>aj</i> <i>uj</i> <i>uj</i> <i>ij</i> <i>uj</i> <i>lji</i> <i>nji</i>	<i>avavši</i> <i>ivavši</i> <i>ovavši</i> <i>ivši</i> <i>uvši</i> <i>avši</i> <i>evši</i>

Flexionsklasse	Infinitiv	Präsens 3. Pl.	Präteritum Sg. feminin	Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpartizip Präteritum
7	<i>psti</i>	<i>bu</i>	<i>bla</i>	<i>be</i>	<i>bah</i>	<i>bi</i>	<i>bavši</i>
	<i>sti</i>	<i>du</i>	<i>la</i>	<i>de</i>	<i>dah</i>	<i>di</i>	<i>davši</i>
	<i>sti</i>	<i>dnu</i>	<i>la</i>	<i>de</i>	<i>dah</i>	<i>dni</i>	<i>vši</i>
	<i>psti</i>	<i>pu</i>	<i>pla</i>	<i>pe</i>	<i>pah</i>	<i>pi</i>	<i>pavši</i>
	<i>sti</i>	<i>su</i>	<i>sla</i>	<i>se</i>	<i>sah</i>	<i>si</i>	<i>savši</i>
	<i>sti</i>	<i>tu</i>	<i>tla</i>	<i>te</i>	<i>tah</i>	<i>ti</i>	<i>tavši</i>
	<i>sti</i>	<i>stu</i>	<i>sla</i>	<i>ste</i>	<i>stah</i>	<i>sti</i>	<i>stavši</i>
	<i>sti</i>	<i>tnu</i>	<i>la</i>	<i>te</i>	<i>tah</i>	<i>tni</i>	<i>vši</i>
	<i>sti</i>	<i>zu</i>	<i>zla</i>	<i>ze</i>	<i>zah</i>	<i>zi</i>	<i>zavši</i>
8	<i>ći</i>	<i>ku</i>	<i>kla</i>	<i>če</i>	<i>čah</i>	<i>ci</i>	<i>kavši</i>
	<i>ći</i>	<i>đu</i>	<i>šla</i>	<i>đe</i>	<i>đah</i>	<i>đi</i>	<i>šavši</i>
	<i>ći</i>	<i>hu</i>	<i>hla</i>	<i>še</i>	<i>šah</i>	<i>si</i>	<i>havši</i>
	<i>ći</i>	<i>gu</i>	<i>gla</i>	<i>že</i>	<i>žah</i>	<i>zi</i>	<i>gavši</i>
	<i>ći</i>	<i>gnu</i>	<i>gla</i>	<i>že</i>	<i>žah</i>	<i>zi</i>	<i>gavši</i>
	<i>ći</i>	<i>du</i>	<i>šla</i>	<i>de</i>	<i>dah</i>	<i>di</i>	<i>šavši</i>
	<i>ći</i>	<i>knu</i>	<i>kla</i>	<i>če</i>	<i>čah</i>	<i>ci</i>	<i>kavši</i>

Abbildung III.6: Tabelle der Flexionsklassen

Wie in Abbildung III.6 zu erkennen ist, können allein durch die Zuordnung eines Verbs zu einer Flexionsklasse die einzelnen Flexionsformen noch nicht abgeleitet werden, da die einzelnen Flexionsklassen viele Unterklassen umfassen. Aus diesem Grund verfeinern wir in unserer Untersuchung die Flexionsklassen mit eigens definierten Flexionstypen.

Für die Codierung der Flexionstypen wird eine Kombination aus artifizierlicher und mnemotechnischer Codierung (siehe II.1.5.2) verwendet. Der artifizierliche Teil besteht aus einer Ziffer am Anfang der Bezeichnung eines Flexionstyps, die für die in Abbildung III.6 definierte Flexionsklasse steht. Unterschiede innerhalb von Flexionsklassen werden mnemotechnisch durch Buchstaben im

Anschluss an die Ziffer bezeichnet. Bei diesen Buchstaben handelt es sich um die Infinitivausgänge der zugehörigen Verben ohne *i*. Verben der Flexionstypen **2at**, **2it** und **2et** gehören zur Flexionsklasse **2**, weisen aber unterschiedliche Infinitivendungen und kleine Unterschiede in der Konjugation auf. Diese Codierung ermöglicht eine eindeutige Kurzbezeichnung jedes Flexionstyps. Falls zwei Flexionstypen bei dieser Codierung dieselbe Bezeichnung erhalten würden, wird sie, durch einen Bindestrich getrennt, um einen Stammausgang erweitert, damit die Eindeutigkeit gewährleistet ist. In den meisten Fällen handelt es sich hierbei um den Präsensstammausgang. Wenn auch dieser bei zwei Flexionstypen identisch ist, wird derjenige Stammausgang genannt, in dem sich die beiden betreffenden Flexionstypen unterscheiden.

Beispiel: **8ć-k**, **8ć-g**, **8ć-h**.

1.5.2.1 Generierungsschema

Da die manuelle Eingabe der Schlüsselformen für jedes der über 4.000 Verben zu mühsam wäre, wird ein Generierungsschema verwendet, mit dessen Hilfe die Schlüsselformen aufgrund von Infinitiv und zugehörigem Flexionstyp erzeugt werden.

In Abbildung III.7 sind die Generierungsregeln in Abhängigkeit der Flexionstypen aufgeführt. Zum besseren Verständnis unseres Generierungsschemas geben wir Erklärungen zu den einzelnen Spalten:

Flexionstyp: In dieser Spalte sind alle verwendeten Flexionstypen aufgelistet.

Die folgenden Spalten beinhalten Angaben zur Ermittlung der Schlüsselformen. Jede teilt sich in die folgenden beiden Unterspalten:

St. (Stelle): In dieser Unterspalte steht die Anzahl der Buchstaben, die im Infinitiv von rechts weggestrichen werden müssen, bevor der Ausgang der Schlüsselform angehängt wird.

Ausg. (Ausgang): In der zweiten Unterspalte steht der neue Ausgang, der an den verbleibenden Rest des Infinitivs angehängt wird. Falls in dieser Spalte nichts steht, wird nach dem Wegstreichen nichts hinzugefügt.

Beispiel: Das Verb **igrati** ist dem Flexionstyp *1at* zugeordnet. Laut Abbildung III.7 werden die Schlüsselformen folgendermaßen gebildet:

Präsens 3. Pl. wird durch Wegstreichen der letzten **drei** (Spalte "Stelle") Buchstaben des Infinitivs und durch Hinzufügen von **aju** (Spalte "Ausgang") gebildet: **igraju**.

Partizip Präteritum Sg. feminin wird durch Wegstreichen der letzten **zwei** (Spalte "Stelle") Buchstaben des Infinitivs und durch Hinzufügen von **la** (Spalte "Ausgang") gebildet: **igrala**.

Aorist 3. Sg. wird durch Wegstreichen der letzten **zwei** (Spalte "Stelle") Buchstaben des Infinitivs und durch Hinzufügen von **∅** (Spalte "Ausgang") gebildet: **igra**.

Imperfekt 1. Sg. wird durch Wegstreichen der letzten **zwei** (Spalte "Stelle") Buchstaben des Infinitivs und durch Hinzufügen von **h** (Spalte "Ausgang") gebildet: **igrah**.

Imperativ Sg. wird durch Wegstreichen der letzten **zwei** (Spalte "Stelle") Buchstaben des Infinitivs und durch Hinzufügen von **j** (Spalte "Ausgang") gebildet: **igraj**.

Adverbialpartizip Präteritum wird durch Wegstreichen der letzten **zwei** (Spalte "Stelle") Buchstaben des Infinitivs und durch Hinzufügen von **vši** (Spalte "Ausgang") gebildet: **igravši**.

Wenn weder in der Spalte „Stelle“ noch in der Spalte „Ausgang“ ein Eintrag vorhanden ist, bedeutet das, dass diese Form für die Verben dieses Flexionstyps grundsätzlich nicht existiert. So ist der Aorist nur bei perfektiven und das Imperfekt nur bei imperfektiven Verben vorhanden. Simplicia sind meistens imperfektiv und werden durch das Voranstellen einer Präposition perfektiviert. Aorist- und Imperfektausgang werden auch genannt, wenn nicht feststellbar ist, ob ein Flexionstyp ausschließlich perfektive oder imperfektive Verben ent-

hält und somit Aorist oder Imperfekt für diesen Flexionstyp tatsächlich vorhanden sind.

Beispielverben zu den einzelnen Flexionstypen finden sich in 3.3.2.

Flexionstyp	Generierungsschema											
	Präsens 3. Pl.		Partizip Präteritum Sg. feminin		Aorist 3. Sg.		Imperfekt 1. Sg.		Imperativ Sg.		Adverbial- partizip Präteritum	
	St.	Ausg.	St.	Ausg.	St.	Ausg.	St.	Ausg.	St.	Ausg.	St.	Ausg.
1at	3	aju	2	la	2	-	2	h	2	j	2	vši
2it	3	e	2	la	2	-	3	ah	3	i	2	vši
2bitmitpitvit	3	e	2	la	2	-	3	ljah	3	i	2	vši
2citkit	3	e	2	la	2	-	4	čah	3	i	2	vši
2dit	3	e	2	la	2	-	4	đah	3	i	2	vši
2gizit	3	e	2	la	2	-	4	žah	3	i	2	vši
2hitsit	3	e	2	la	2	-	4	šah	3	i	2	vši
2lit	3	e	2	la	2	-	4	ljah	3	i	2	vši
2nit	3	e	2	la	2	-	4	njah	3	i	2	vši
2slit	3	e	2	la	2	-	5	šljah	3	i	2	vši
2tit	3	e	2	la	2	-	4	čah	3	i	2	vši
2tit-šč	3	e	2	la	2	-	5	ščah	3	i	2	vši
2tit-št	3	e	2	la	2	-	5	štah	3	i	2	vši
2znit	3	e	2	la	2	-	5	žnjah	3	i	2	vši
2snit	3	e	2	la	2	-	5	šnjah	3	i	2	vši
2et	3	e	2	la	2	-			3	i	2	vši
2at	3	e	2	la	2	-	2	h	3	i	2	vši
3at	3	u	2	la	2	-	2	h	3	i	2	vši
3et-n	3	nu	2	la	2	-			3	ni	2	vši
3catkat	4	ču	2	la	2	-	2	h	4	či	2	vši
3dat	4	đu	2	la	2	-	2	h	4	đi	2	vši
3gatzat	4	žu	2	la	2	-	2	h	4	ži	2	vši
3hatsat	4	šu	2	la	2	-	2	h	4	ši	2	vši
3tat	4	ču	2	la	2	-	2	h	4	ći	2	vši
3ijet	5	esu	2	la	2	-	3	ah	5	esi	2	vši
3jat	3	u	2	la	2	-	2	h	3	-	2	vši
3klet	4	unu	2	la	2	-			4	uni	2	vši
3et-m	3	mu	2	la	2	-			3	mi	2	vši
3ut-m	3	mu	2	la	2	-			3	mi	2	vši
3at-n	2	nu	2	la	2	-	2	h	2	ni	2	vši
3ut-p	3	pu	2	la	2	-			3	pi	2	vši
3rat	4	eru	2	la	2	-	2	h	4	eri	2	vši
3rijet	5	u	5	la	2	-	5	ah	4	-	2	vši
3ut	3	u	2	la	2	-	3		3	i	2	vši
3vat	4	ovu	2	la	2	-	2	h	4	ovi	2	vši
4bjetmjetpjetvjet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši

Flexionstyp	Generierungsschema											
	Präsens 3. Pl.		Partizip Präteritum Sg. feminin		Aorist 3. Sg.		Imperfekt 1. Sg.		Imperativ Sg.		Adverbial- partizip Präteritum	
	St.	Ausg.	St.	Ausg.	St.	Ausg.	St.	Ausg.	St.	Ausg.	St.	Ausg.
4cjetkjet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši
4djet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši
4gjetzjet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši
4hjetsjet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši
4jet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši
4ljet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši
4njet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši
4sljet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši
4tjet	4	iju	2	la	2	-	4	ijah	4	ij	2	vši
5bjetmjetpjetvjet	4	e	2	la	2	-	4	ljah	4	i	2	vši
5cjetkjet	4	e	2	la	2	-	5	čah	4	i	2	vši
5djet	4	e	2	la	2	-	5	đah	4	i	2	vši
5gjetzjet	4	e	2	la	2	-	5	žah	4	i	2	vši
5hjetsjet	4	e	2	la	2	-	5	šah	4	i	2	vši
5jet	4	e	2	la	2	-	4	ah	4	i	2	vši
5ljet	4	e	2	la	2	-	5	ljah	4	i	2	vši
5njet	4	e	2	la	2	-	5	njah	4	i	2	vši
5sljet	4	e	2	la	2	-	5	šljah	4	i	2	vši
5tjet	4	e	2	la	2	-	5	ćah	4	i	2	vši
6avat	4	ju	2	la	2	-	2	h	2	j	2	vši
6ivatovat	5	uju	2	la	2	-	2	h	5	uj	2	vši
6itut	2	ju	2	la	2	-	2	jah	2	j	2	vši
6klat	4	olju	2	la	2	-	2	h	4	olji	2	vši
6at	3	lju	2	la	2	-	2	h	3	lji	2	vši
6ljet	5	elju	2	la	2	-	2	h	5	elji	2	vši
6et	3	nju	2	la	2	-			3	nji	2	vši
6slat	5	šalju	2	la	2	-	2	h	5	šalji	2	vši
7st-b	4	bu	4	bla	4	be	4	bah	4	bi	4	bavši
7st-d	3	du	3	la	3	de	3	dah	3	di	3	davši
7st-od	3	du	3	la	3	de	3	dah	3	di	3	davši
7st-dn	3	dnu	3	la	3	de	3	dah	3	dni	3	vši
7st-p	3	u	3	la	3	e	3	ah	3	i	3	avši
7st-s	2	u	2	la	2	e	2	ah	2	i	2	avši
7st-t	3	tu	3	la	3	te	3	tah	3	ti	3	tavši
7st-st	1	u	2	la	1	e	1	ah	0	-	1	avši
7st-tn	3	tnu	3	la	3	te	3	tah	3	tni	3	vši
7st-z	3	zu	3	zla	3	ze	3	zah	3	zi	3	zavši
8ć-k	2	ku	2	kla	2	če	2	čah	2	ci	2	kavši
8ć-đ	2	đu	2	šla	2	đe	2	đah	2	đi	2	šavši
8ć-h	2	hu	2	hla	2	še	2	šah	2	si	2	havši
8ć-g	2	gu	2	gla	2	že	2	žah	2	zi	2	gavši

Flexionstyp	Generierungsschema											
	Präsens 3. Pl.		Partizip Präteritum Sg. feminin		Aorist 3. Sg.		Imperfekt 1. Sg.		Imperativ Sg.		Adverbial- partizip Präteritum	
	St.	Ausg.	St.	Ausg.	St.	Ausg.	St.	Ausg.	St.	Ausg.	St.	Ausg.
8ć-gn	2	gnu	2	gla	2	že	2	žah	2	zi	2	gavši
8ć-d	2	du	2	šla	2	de	2	dah	2	di	2	šavši
8ć-kn	2	knu	2	kla	2	če	2	čah	2	ci	2	kavši

Abbildung III.7: Tabelle der Flexionstypen

1.6 Ableitungsregeln und Ausnahmeformen

In diesem Abschnitt werden die Ableitungsregeln für die Nichtschlüssel­formen (1.6.1) der kroatischen Verben mit ihren Ausnahmen (1.6.2) spezifiziert. Eine Darstellung wie in den gängigen Grammatiken wird nicht verwendet.

1.6.1 Die Ableitungsregeln

Die nachfolgenden Ableitungsregeln beschreiben die Herleitung der Nicht­Schlüssel­formen aus den Schlüssel­formen. Eine Übersicht über die Flexions­formen und die für deren Herleitung jeweils notwendige Schlüssel­form geben Abbildung III.8 und Abbildung III.9. Die Grafik in Abbildung III.8 zeigt die zur Herleitung der Flexions­formen verwendete Stammverteilung. Jede Schraffur steht für eine der Schlüssel­formen.



vom Infinitiv abzuleitende Formen



vom Präsens 3. Pl. abzuleitende Formen



vom Partizip Präteritum Sg. feminin abzuleitende Formen



vom Aorist 3. Sg. abzuleitende Formen



vom Imperfekt 1. Sg. abzuleitende Formen



vom Imperativ Sg. abzuleitende Formen



vom Adverbialpartizip Präteritum abzuleitende Formen

So ist z.B. Präsens 2. Sg. abgeleitet von Präsens 3. Pl., zu erkennen an der gleichen Schraffur ().

Die Tabelle in Abbildung III.9 gibt einen Überblick über die Zuordnung der Flexionsformen zu den Schlüsselformen und evtl. notwendige Zwischenformen (cf. II.1.6.1).

Infinitiv	
-----------	--

Präsens

1. Sg.	
2. Sg.	
3. Sg.	
1. Pl.	
2. Pl.	
3. Pl.	

Partizip Präteritum

Sg. m.	
Sg. f.	
Sg. n.	
Pl. m.	
Pl. f.	
Pl. n.	

Aorist

	Infinitiv endet auf <i>-sti</i> oder <i>-ći</i>	Infinitiv endet nicht auf <i>-sti</i> oder <i>-ći</i>
1. Sg.		
2. Sg.		
3. Sg.		
1. Pl.		
2. Pl.		
3. Pl.		

Imperfekt

1. Sg.	
2. Sg.	
3. Sg.	
1. Pl.	
2. Pl.	
3. Pl.	

Imperativ

2. Sg.	
1. Pl.	
2. Pl.	

Adverbialpartizipien

Präs.	
Prät.	

Partizip Passiv

sonst	Infinitiv auf <i>-eti</i> (nicht auf <i>-jeti</i>), <i>-iti</i> (mit Präs. 3. Pl. auf <i>-iju</i>), <i>-uti</i>

Abbildung III.8: Stammverteilung

	Flexionsform	Zwischenform	Schlüsselform
	Infinitiv	--	Infinitiv
Präsens (1.6.1.1 und 1.6.1.2)	Präsens 1. Sg.	--	Präsens 3. Pl.
	Präsens 2. Sg.	Präsens 1. Sg.	Präsens 3. Pl.
	Präsens 3. Sg.	Präsens 1. Sg.	Präsens 3. Pl.
	Präsens 1. Pl.	Präsens 1. Sg.	Präsens 3. Pl.
	Präsens 2. Pl.	Präsens 1. Sg.	Präsens 3. Pl.
	Präsens 3. Pl.	--	Präsens 3. Pl.
Partizip- Präteritum (1.6.1.3)	Partizip Präteri- tum Sg. maskulin	--	Partizip Präteri- tum Sg. feminin
	Partizip Präteri- tum Sg. feminin	--	Partizip Präteri- tum Sg. feminin
	Partizip Präteri- tum Sg. neutrum	--	Partizip Präteri- tum Sg. feminin
	Partizip Präteri- tum Pl. maskulin	--	Partizip Präteri- tum Sg. feminin
	Partizip Präteri- tum Pl. feminin	--	Partizip Präteri- tum Sg. feminin
	Partizip Präteri- tum Pl. neutrum	--	Partizip Präteri- tum Sg. feminin
Aorist (1.6.1.4)	Aorist 1. Sg.	Lexemstamm	Infinitiv und Präsens 3. Pl
	Aorist 2. Sg.	--	Aorist 3. Sg.
	Aorist 3. Sg.	--	Aorist 3. Sg.
	Aorist 1. Pl.	Lexemstamm	Infinitiv und Präsens 3. Pl
	Aorist 2. Pl.	Lexemstamm	Infinitiv und Präsens 3. Pl
	Aorist 3. Pl.	Lexemstamm	Infinitiv und Präsens 3. Pl

	Flexionsform	Zwischenform	Schlüsselform
Imperfekt (1.6.1.5)	Imperfekt 1. Sg.	--	Imperfekt 1. Sg.
	Imperfekt 2. Sg.	--	Imperfekt 1. Sg.
	Imperfekt 3. Sg.	--	Imperfekt 1. Sg.
	Imperfekt 1. Pl.	--	Imperfekt 1. Sg.
	Imperfekt 2. Pl.	--	Imperfekt 1. Sg.
	Imperfekt 3. Pl.	--	Imperfekt 1. Sg.
Imperativ (1.6.1.6)	Imperativ 2. Sg.	--	Imperativ 2. Sg.
	Imperativ 1. Pl.	--	Imperativ 2. Sg.
	Imperativ 2. Pl.	--	Imperativ 2. Sg.
Adverbialpartizip (1.6.1.7)	Adverbialpartizip Präsens	--	Präsens 3. Pl.
	Adverbialpartizip Präteritum	--	Adverbialpartizip Präteritum
Partizip Passiv (1.6.1.8)	Partizip Passiv	--	Imperfekt 1. Sg. und Infinitiv

Abbildung III.9: Übersicht der Flexionsformen

Auf der Basis der obigen Tabelle ergeben sich folgende Ableitungsregeln. Nach jeder Gruppe von Flexionsformen werden Beispiele genannt. Sie werden folgendermaßen dargestellt:

Ergebnis der Konkatenation ***Ergebnis der Dekatenation*** **Ausgangsform** (*Infinitiv*)

1.6.1.1 Präsens 1. Sg.

De-/Konkatenation	Bedingung
Präsens 1. Sg. = Präsens 3. Pl. ⊖ <i>gu</i> ⊕ <i>žem</i>	Präsens 3. Pl. endet auf <i>-gu</i>
Präsens 1. Sg. = Präsens 3. Pl. ⊖ <i>ku</i> ⊕ <i>čem</i>	Präsens 3. Pl. endet auf <i>-ku</i>
Präsens 1. Sg. = Präsens 3. Pl. ⊖ <i>hu</i> ⊕ <i>šem</i>	Präsens 3. Pl. endet auf <i>-hu</i>
Präsens 1. Sg. = Präsens 3. Pl. ⊖ <i>ju</i> ⊕ <i>m</i>	Präsens 3. Pl. endet auf <i>-aju</i>
Präsens 1. Sg. = Präsens 3. Pl. ⊖ <i>u</i> ⊕ <i>em</i>	Präsens 3. Pl. endet auf <i>-u</i> und Präsens 3. Pl. endet <u>nicht</u> auf <i>-gu, -ku, -hu</i> oder <i>-aju</i>
Präsens 1. Sg. = Präsens 3. Pl. ⊖ <i>e</i> ⊕ <i>im</i>	Präsens 3. Pl. endet auf <i>-e</i>

Beispiele:

<i>strižem</i>	<i>stri-</i>	<u><i>strigu</i></u>	(<i>strići</i>)
<i>pečem</i>	<i>pe-</i>	<u><i>peku</i></u>	(<i>peći</i>)
<i>vršem</i>	<i>vr-</i>	<u><i>vrhu</i></u>	(<i>vrći</i>)
<i>igram</i>	<i>igra-</i>	<u><i>igraju</i></u>	(<i>igrati</i>)
<i>krenem</i>	<i>kren-</i>	<u><i>krenu</i></u>	(<i>krenuti</i>)
<i>vidim</i>	<i>vid-</i>	<u><i>vide</i></u>	(<i>vidjeti</i>)

Ausnahmen:

<i>-dajem</i>	<i>-daj-</i>	<u><i>-daju</i></u>	(<i>-davati</i>)
<i>-znajem</i>	<i>-znaj-</i>	<u><i>-znaju</i></u>	(<i>-znavati</i>)
<i>jesam</i>	<i>jes-</i>	<u><i>jesu</i></u>	(<i>biti</i>)
<i>hoću</i>	<i>hoć-</i>	<u><i>hoće</i></u>	(<i>htjeti</i>)
<i>mogu</i>	<i>mog-</i>	<u><i>mogu</i></u>	(<i>moći</i>)

1.6.1.2 Präsens 2. Sg. bis 2. Pl.

De-/Konkatenation	Bedingung
Präsens 2. Sg. = Präsens 1. Sg. \ominus <i>m</i> \oplus <i>š</i>	alle

Beispiele:

igraš *igra-* *igram* (*igrati*)
vidiš *vidi-* *vidim* (*vidjeti*)

Ausnahmen:

jesi *jesam* (*biti*)
hoćeš *hoću* (*htjeti*)
možeš *možu* (*moći*)

De-/Konkatenation	Bedingung
Präsens 3. Sg. = Präsens 1. Sg. \ominus <i>m</i> \oplus \emptyset	alle

Beispiele:

igra *igra-* *igram* (*igrati*)
vidi *vidi-* *vidim* (*vidjeti*)

Ausnahmen:

jest *jesam* (*biti*)
hoće *hoću* (*htjeti*)
može *možu* (*moći*)

De-/Konkatenation	Bedingung
Präsens 1. Pl. = Präsens 1. Sg. \ominus <i>m</i> \oplus <i>mo</i>	alle

Beispiele:

igramo ***igra-*** **igram** (*igrati*)
vidimo ***vidi-*** **vidim** (*vidjeti*)

Ausnahmen:

jesmo **jesam** (*biti*)
hoćemo **hoću** (*htjeti*)
možemo **moqu** (*moći*)

De-/Konkatenation	Bedingung
Präsens 2. Pl. = Präsens 1. Sg. \ominus <i>m</i> \oplus <i>te</i>	alle

Beispiele:

igrate ***igra-*** **igram** (*igrati*)
vidite ***vidi-*** **vidim** (*vidjeti*)

Ausnahmen:

jeste **jesam** (*biti*)
hoćete **hoću** (*htjeti*)
možete **moqu** (*moći*)

1.6.1.3 Partizip Präteritum

De-/Konkatenation	Bedingung
Partizip Präteritum Sg. maskulin = Partizip Präteritum Sg. feminin ⊖ <i>la</i> ⊕ <i>ao</i>	Infinitiv endet auf <i>-ći</i> oder <i>-sti</i> und Partizip Präteritum Sg. feminin hat vor der Endung <i>-la</i> einen <u>Konsonantbuchstaben</u>
Partizip Präteritum Sg. maskulin = Partizip Präteritum Sg. feminin ⊖ <i>jela</i> ⊕ <i>io</i>	Infinitiv endet auf <i>-jeti</i>
Partizip Präteritum Sg. maskulin = Partizip Präteritum Sg. feminin ⊖ <i>la</i> ⊕ <i>o</i>	sonst

Beispiele:

<i>išao</i>	<i>iš-</i>	<u><i>išla</i></u>	(<i>ići</i>)
<i>vidio</i>	<i>vid-</i>	<u><i>vidjela</i></u>	(<i>vidjeti</i>)
<i>igrao</i>	<i>igra-</i>	<u><i>igrala</i></u>	(<i>igrati</i>)

De-/Konkatenation	Bedingung
Partizip Präteritum Sg. neutrum = Partizip Präteritum Sg. feminin ⊖ <i>a</i> ⊕ <i>o</i>	alle

Beispiele:

<i>vidjelo</i>	<i>vidjel-</i>	<u><i>vidjela</i></u>	(<i>vidjeti</i>)
<i>igralo</i>	<i>igral-</i>	<u><i>igrala</i></u>	(<i>igrati</i>)

De-/Konkatenation	Bedingung
Partizip Präteritum Pl. maskulin = Partizip Präteritum Sg. feminin ⊖ a ⊕ i	alle
Partizip Präteritum Pl. feminin = Partizip Präteritum Sg. feminin ⊖ a ⊕ e	alle
Partizip Präteritum Pl. neutrum = Partizip Präteritum Sg. feminin	alle

Beispiele:

<i>vidjeli</i>	vidjel-	<u>vidjela</u>	(<i>vidjeti</i>)
<i>igrali</i>	igral-	<u>igrala</u>	(<i>igrati</i>)
<i>vidjele</i>	vidjel-	<u>vidjela</u>	(<i>vidjeti</i>)

1.6.1.4 Aorist und Lexemstamm

Zur Bildung einiger Aoristformen wird der Lexemstamm benötigt, dessen Ableitungsregeln vorangestellt werden.

De-/Konkatenation	Bedingung
Lexemstamm = Präsens 3. Pl. ⊖ nu	Infinitiv endet auf <i>-sti</i> oder <i>-ći</i> und Präsens 3. Pl. endet auf <i>-nu</i>
Lexemstamm = Präsens 3. Pl. ⊖ u	Infinitiv endet auf <i>-sti</i> oder <i>-ći</i> und Präsens 3. Pl. endet <u>nicht</u> auf <i>-nu</i>
Lexemstamm = Infinitiv ⊖ ti	Infinitiv endet nicht auf <i>-sti</i> oder <i>-ći</i>

Beispiele:

rek- **reknu** **reći**
 id- **idu** **ići**
 igra- **igrati**

Auf dem ermittelten Lexemstamm basieren folgende Ableitungsregeln.

De-/Konkatenation	Bedingung
Aorist 1. Sg. = Lexemstamm ⊕ h	Lexemstamm endet auf einen <u>Vokalbuchstaben</u>
Aorist 2. Sg. = Aorist 3. Sg.	
Aorist 3. Sg. = Schlüsselwort	
Aorist 1. Pl. = Lexemstamm ⊕ sno	
Aorist 2. Pl. = Lexemstamm ⊕ ste	
Aorist 3. Pl. = Lexemstamm ⊕ še	

Beispiele:

pomislih **pomisli-** (*pomisliti*)
naigrah **naigra-** (*naigrati*)

pomislismo **pomisli-** (*pomisliti*)
naigrasmo **naigra-** (*naigrati*)

pomisliše **pomisli-** (*pomisliti*)
naigraše **naigra-** (*naigrati*)

De-/Konkatenation	Bedingung
Aorist 1. Sg. = Lexemstamm ⊕ <i>oh</i>	Lexemstamm endet auf einen <u>Konsonantbuchstaben</u>
Aorist 2. Sg. = Aorist 3. Sg.	
Aorist 3. Sg. = Schlüsselform	
Aorist 1. Pl. = Lexemstamm ⊕ <i>osmo</i>	
Aorist 2. Pl. = Lexemstamm ⊕ <i>oste</i>	
Aorist 3. Pl. = Lexemstamm ⊕ <i>oše</i>	

Beispiele:

<i>rekoh</i>	<u>rek-</u>	(<i>reći</i>)
<i>strigoh</i>	<u>strig-</u>	(<i>strići</i>)
<i>rekosmo</i>	<u>rek-</u>	(<i>reći</i>)
<i>strigosmo</i>	<u>strig-</u>	(<i>strići</i>)
<i>rekoše</i>	<u>rek-</u>	(<i>reći</i>)
<i>strigoše</i>	<u>strig-</u>	(<i>strići</i>)

1.6.1.5 Imperfekt

De-/Konkatenation	Bedingung
Imperfekt 1. Sg. = Schlüsselform	Alle
Imperfekt 2. Sg. = Imperfekt 1. Sg. \ominus <i>h</i> \oplus <i>še</i>	
Imperfekt 3. Sg. = Imperfekt 1. Sg. \ominus <i>h</i> \oplus <i>še</i>	
Imperfekt 1. Pl. = Imperfekt 1. Sg. \ominus <i>h</i> \oplus <i>sno</i>	
Imperfekt 2. Pl. = Imperfekt 1. Sg. \ominus <i>h</i> \oplus <i>ste</i>	
Imperfekt 3. Pl. = Imperfekt 1. Sg. \oplus <i>u</i>	

Beispiele:

<i>igraše</i>	igra-	<u>igrah</u>	(<i>igrati</i>)
<i>nošaše</i>	noša-	<u>nošah</u>	(<i>nositi</i>)
<i>igrasno</i>	igra-	<u>igrah</u>	(<i>igrati</i>)
<i>nošasno</i>	noša-	<u>nošah</u>	(<i>nositi</i>)
<i>igraste</i>	igra-	<u>igrah</u>	(<i>igrati</i>)
<i>nošaste</i>	noša-	<u>nošah</u>	(<i>nositi</i>)
<i>igrahu</i>		<u>igrah</u>	(<i>igrati</i>)
<i>nošahu</i>		<u>nošah</u>	(<i>nositi</i>)

1.6.1.6 Imperativ

De-/Konkatenation	Bedingung
Imperativ 1. Pl. = Imperativ 2. Sg. ⊕ <i>mo</i>	Alle

Beispiele:

igrajmo *igraj* (*igrati*)
nosimo *nosi* (*nositi*)

De-/Konkatenation	Bedingung
Imperativ 2. Pl. = Imperativ 2. Sg. ⊕ <i>te</i>	Alle

Beispiele:

igrajte *igraj* (*igrati*)
nosite *nosi* (*nositi*)

1.6.1.7 Adverbialpartizip Präsens

Bemerkung zum Vorkommen: Diese Flexionsform „kann nur von imperfektiven Verben gebildet werden“ (Engel; Mrazović 1986: 109).

De-/Konkatenation	Bedingung
Adverbialpartizip Präsens = Präsens 3. Pl. ⊕ <i>ći</i>	Alle

Beispiele:

igrajući *igraju* (*igrati*)
noseći *nose* (*nositi*)

Ausnahmen:

budući *jesu* (*biti*)
hoteći *hoće* (*htjeti*)

1.6.1.8 Partizip Passiv

Das Partizip Passiv „wird nur von Verben mit Akkusativerganzung (»transitiven Verben«) gebildet“ (Engel; Mrazovic 1986: 111). Die folgenden Ableitungsregeln basieren auf der Alternationstabelle nach Kunzmann-Muller 2002: 34-35.

De-/Konkatenation	Bedingung
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>ah</i> ⊕ <i>en</i>	Infinitiv endet auf <i>-ci</i> , <i>-sti</i> oder <i>-iti</i> und Prasens 3. Pl. endet <u>nicht</u> auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>ah</i> ⊕ <i>en</i>	Infinitiv endet auf <i>-jeti</i> und Prasens 3. Pl. endet auf <i>-e</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>ijah</i> ⊕ <i>ljen</i>	Infinitiv endet auf <i>-bjeti</i> , <i>-mjeti</i> , <i>-pjeti</i> oder <i>-vjeti</i> und Prasens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>cijah</i> ⊕ <i>cen</i>	Infinitiv endet auf <i>-cjeti</i> und Prasens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>kijah</i> ⊕ <i>cen</i>	Infinitiv endet auf <i>-kjeti</i> und Prasens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>dijah</i> ⊕ <i>den</i>	Infinitiv endet auf <i>-djeti</i> und Prasens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>gijah</i> ⊕ <i>zen</i>	Infinitiv endet auf <i>-gjeti</i> und Prasens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>zijah</i> ⊕ <i>zen</i>	Infinitiv endet auf <i>-zjeti</i> und Prasens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>hijah</i> ⊕ <i>šen</i>	Infinitiv endet auf <i>-hjeti</i> und Prasens 3. Pl. endet auf <i>-iju</i>

De-/Konkatenation	Bedingung
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>sijah</i> ⊕ <i>šen</i>	Infinitiv endet auf <i>-sjeti</i> und Präsens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>ijah</i> ⊕ <i>jen</i>	Infinitiv endet auf <i>-ljeti</i> oder <i>-njeti</i> und Präsens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>slijah</i> ⊕ <i>šljen</i>	Infinitiv endet auf <i>-sljeti</i> und Präsens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>tijah</i> ⊕ <i>ćen</i>	Infinitiv endet auf <i>-tjeti</i> und Präsens 3. Pl. endet auf <i>-iju</i>
Partizip Passiv = Imperfekt 1. Sg. ⊖ <i>h</i> ⊕ <i>n</i>	Infinitiv endet auf <i>-ati</i>
Partizip Passiv = Infinitiv ⊖ <i>i</i>	Infinitiv endet auf <i>-eti</i> oder <i>-uti</i> und Infinitiv endet <u>nicht</u> auf <i>-jeti</i>
Partizip Passiv = Infinitiv ⊖ <i>i</i>	Infinitiv endet auf <i>-iti</i> und Präsens 3. Pl. endet auf <i>-iju</i>

Beispiele:

<i>strižen</i>	striž-	<u>strižah</u>	(<i>strići</i>)
<i>viđen</i>	viđ-	<u>viđah</u>	(<i>vidjeti</i>)
<i>razumljen</i>	razum-	<u>razumijah</u>	(<i>razumjeti</i>)
<i>igran</i>	igra-	<u>igrah</u>	(<i>igrati</i>)
<i>raspet</i>		<u>raspeti</u>	
<i>pit</i>		<u>piti</u>	

1.6.2 Ausnahmeformen

Hier werden diejenigen Ausnahmen systematisch zusammengestellt, die in vorherigen Abschnitten erwähnt wurden. Allerdings wird absolute Vollständigkeit im Hinblick auf alle überhaupt vorkommenden Ausnahmen nicht angestrebt und nicht erreicht.

Infinitiv	Präsens				
	1. Sg.	2. Sg.	3. Sg.	1. Sg.	2. Pl.
<i>moći</i>	<i>mogu</i>	<i>možeš</i>	<i>može</i>	<i>možemo</i>	<i>možete</i>
<i>-davati</i>	<i>-dajem</i>	<i>-daješ</i>	<i>-daje</i>	<i>-dajemo</i>	<i>-dajete</i>
<i>-znavati</i>	<i>-znajem</i>	<i>-znaješ</i>	<i>-znaje</i>	<i>-znajemo</i>	<i>-znajete</i>
<i>htjeti</i>	<i>hoću</i>	<i>hoćeš</i>	<i>hoće</i>	<i>hoćemo.</i>	<i>hoćete</i>
<i>biti</i>	<i>jesam</i>	<i>jesi</i>	<i>jest</i>	<i>jesmo</i>	<i>jeste</i>

Infinitiv	Adverbialpartizip Präsens
<i>biti</i>	<i>budući</i>
<i>htjeti</i>	<i>hoteći</i>

1.7 Lexeme mit zwei Flexionstypen

Verben mit mehreren Konjugationen, die sich semantisch voneinander unterscheiden, werden als mehrere Lexeme interpretiert (cf. II.1.7.1). Diese werden in 4 gesondert aufgeführt, wobei jedem Verb der entsprechende Flexionstyp zugeordnet wird.

Infinitiv	Bedeutung	Flexionstyp
<i>1pasti</i>	stürzen	7st-dn
<i>2pasti</i>	weiden	7st-s

Verben mit mehreren Konjugationen, die sich semantisch nicht voneinander unterscheiden, werden in 4 nur einmal angegeben, wobei der zugeordnete Flexionstyp die gebräuchlichste Konjugation des Verbs angibt.

2. Durchführung der Datenanalyse

Dieser Abschnitt befasst sich mit der Datenanalyse des sprachlichen Materials speziell für die kroatische Verbalmorphologie. In 2.1 werden die technischen Voraussetzungen und in 2.2 die Anpassungen des Algorithmus beschrieben.

2.1 Technische Voraussetzungen

Für den Aufbau und die Analyse der Datenbasis wurde die Programmiersprache Active Perl in Verbindung mit dem Tabellenkalkulationsprogramm MS Excel verwendet. Für die Nutzung der Dateien auf der beigefügten CD wird aus diesem Grund die Verfügbarkeit von MS Excel und dessen Kenntnis vorausgesetzt.

2.2 Das Data-Mining-Konzept

Für den Algorithmus sind die Anpassungen aus 1.2 zu beachten. Änderungen am Algorithmus speziell für das Kroatische gibt es nicht.

Zusätzlich wird die einzelsprachabhängige Tabelle **Flexionstypen** verwendet. Diese Tabelle enthält alle Konjugationstypen der kroatischen Verben mit den dazugehörigen Ausgängen (cf. 1.5.2.1). Diese Tabelle ist auf der beigefügten CD unter dem Namen **HR_Flexionstypen.xls** zu finden. Der Inhalt dieser Tabelle ist in Abbildung III.7 aufgeführt.

2.2.1 Vorbereitung - Sortieralgorithmus

In Excel werden die Buchstaben *dž*, *lj* und *nj* in Excel als jeweils zwei Buchstaben interpretiert. Dies hat zur Folge, dass sich eine falsche Reihenfolge bei der Sortierung ergeben würde (cf. 1.1). Damit das Verbreregister alphabetisch richtig sortiert wird, werden alle Buchstaben des kroatischen Alphabets mit Hilfe von Excel-Formeln durch alphanumerische Zeichen des Standardzeichensatzes so ersetzt, dass die Wörter beim Sortieren richtig angeordnet werden. So wird beispielsweise dem Digraph *lj*, der ja aus zwei Buchstaben besteht, ein einzelner Buchstabe zugeordnet. Dieser Buchstabe muss sich in der al-

phabetischen Reihenfolge hinter dem zugewiesenen Buchstaben für *l* und vor dem Buchstaben für *m* befinden. Die Reihenfolge in der nachstehenden Tabelle gibt die Sortierreihenfolge wieder.

a	b	c	č	ć	d	dž	đ	e	f	g	h	i	j	k	l
5	6	7	8	9	a	b	c	d	e	f	g	h	i	j	k
k	lj	m	n	nj	o	p	r	s	š	t	u	v	z	ž	
l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	

Abbildung III.10: Substitutionstabelle

2.2.2 Datenstruktur und Algorithmus des Data-Mining-Konzepts

Das Verbrregister setzt sich aus elf Spalten zusammen. In der ersten Spalte werden homogene Cluster und Gemeincluster als solche gekennzeichnet. Verben, die nach der Reduzierung der Verbliste wieder eingefügt wurden (cf. 3.2.1.3), werden mit einem x markiert. Die weiteren Spalten enthalten den Flexionstyp, Aspekt sowie die sieben Schlüsselformen. In der letzten Spalte werden alle Verben aufgelistet, die in einem homogenen Cluster liegen.

3. Nachbereitung der Datenanalyse

Das Data-Mining-Ergebnis wird mit dem Ziel einer besseren Benutzbarkeit aufbereitet. Hierzu zählen typographische Kennzeichnung (3.1) und Reduktionen (3.2). Zusätzlich wird angegeben, welche linguistischen Ergebnisse aus dem Verbregister gewonnen werden können (3.3).

3.1 Typographische Kennzeichnung

Im Verbregister kennzeichnen wir Gemeincluster und homogene Cluster durch Fettdruck (siehe 3.2.1).

3.2 Reduktionen

In diesem Abschnitt werden die Methoden vorgestellt, die zur Reduktion des Verbregisters dienen. Die Motivation zur Reduktion entspricht der in II.3.2. In 3.2.1 werden die formalisierbaren und in 3.2.2 die nicht formalisierbaren Methoden genannt.

3.2.1 Formalisierbare Reduktionen

In diesem Abschnitt werden die formalisierbaren Reduktionen nach II.3.2.1 aufgeführt.

3.2.1.1 Reduktion präfigierter Lexeme in homogenen Clustern

Im Verbregister (4) werden präfigierte Verben weitgehend reduziert. Nur in den folgenden Fällen erscheinen sie:

1. Falls in einem Cluster kein Simplex vorhanden ist, wird jeweils das rückläufig erste von ihnen als sogenanntes „Pseudosimplex“ im Verbregister aufgeführt.
2. Präfigierte Verben, die anders als das jeweilige Simplex (oder „Pseudosimplex“) konjugieren, werden im Verbregister aufgeführt. Nachfolgend

werden alle im Verbbregister aufgelisteten Simplicia genannt, bei denen dieser Fall eintritt.

Simplex („Pseudo-simplex“)	Flexionstyp des Simplex („Pseudo-simplex“)	Präfigiertes Verb (rückläufig erstes)	Flexionstyp der präfigierten Verben	Bedeutung des Simplex („Pseudo-simplex“)
<i>ići</i>	8ć-d	<i>na-ići</i>	8ć-đ	'gehen'
<i>raza-šiljati</i>	3at	<i>iza-šiljati</i>	1at	'verschicken'
<i>kapati</i>	6at	<i>is-kapati</i>	1at	'tropfen'
<i>birati</i>	1at	<i>oda-birati</i>	3at	'wählen'
<i>sisati</i>	1at	<i>i-sisati</i>	3hatsat	'saugen'
<i>mastiti</i>	2tit-št	<i>za-mastiti</i>	2tit-šč	'fetten'

Abbildung III.11: Präfigierte Verben mit anderer Konjugation als ihr Simplex

3.2.1.2 Reduktion innerhalb homogener Cluster

Homogene Cluster werden im Verbbregister zu einem Eintrag zusammengefasst. Die dem homogenen Cluster zugeordneten Verben erscheinen in der Kommentarspalte. Hierbei werden nur Simplicia und Pseudosimplicia aufgeführt. Nur bei den homogenen Clustern *~ovati*, *~čiti*, *~diti*, *~jiti*, *~siti*, *~atiti*, *~ziti*, *~žiti* und *~nuti* werden in der Kommentarspalte keine Verben genannt, da diese Cluster so viele Verben enthalten, dass die Übersichtlichkeit bei deren Nennung leiden würde.

3.2.1.3 Reduktion in Gemeinclustern

Nach dem ersten Durchlauf des Data-Mining-Algorithmus weisen viele nicht benachbarte Cluster jeweils gleiche morphologische Eigenschaften auf. Diese

gehören zu Konjugationstypen mit relativ vielen Verben. Aus diesem Grund werden diese zu von uns so genannten **Gemeinclustern** zusammengefasst. Aus Gründen der Übersichtlichkeit wird deren Disjunktheit verlangt. Jedes Gemeincluster ist daher definiert durch die Kombination aus einem eindeutig bestimmten Infinitivausgang und einem Flexionstyp (cf. II.3.2.1.3). In Abbildung III.12 werden die Gemeincluster aufgeführt, die wir für das kroatische Verbalsystem definieren.

Bezeichnung des Gemeinclusters	Flexionstyp
<i>~ati \ ~ivati</i>	1at
<i>~ivati</i>	6ivatovat
<i>~niti \ ~sniti \ ~zniti</i>	2nit
<i>~riti</i>	2it
<i>~viti</i>	2bitmitpitvit

Abbildung III.12: Gemeincluster im Kroatischen

Gemeincluster werden im Verbregister zu einem einzelnen Eintrag zusammengefasst. Die zu einem Gemeincluster gehörigen Verben werden nicht mehr genannt. Nur im folgenden Fall werden solche Verben aufgeführt. Nach der Reduktion des Verbregisters könnten Gruppen benachbarter Verben vom Leser fälschlicherweise als homogene Cluster interpretiert werden. In diesem Fall müssen Verben, die einem Gemeincluster zugeordnet sind und somit eigentlich nicht aufgeführt werden sollten, wieder eingefügt werden, damit der falsche Anschein nicht entsteht. Solche Verben werden in der Clusterspalte des Verbregisters mit einem x markiert. Dies soll im nachfolgendem Beispiel veranschaulicht werden.

Im Verbbregister sind folgende Verben vorhanden:

6avat	ipf	<i>davati</i>
1at	ipf	<i>sa-vladavati</i>
6avat	ipf	<i>pre-udavati</i>

Da alle Verben, die auf *-ati* enden und den Flexionstyp 1at besitzen, dem Gemeincluster $\sim\text{ati} \setminus \sim\text{ivati}$ zugeordnet sind, würde man das Verb *sa-vladavati* eigentlich nicht aufführen. Dann entstünde jedoch der Eindruck, dass die beiden Verben *davati* und *pre-udavati* ein homogenes Cluster *davati* bilden. Da dies nicht der Fall ist, muss *sa-vladavati* im Verbbregister wieder eingefügt werden.

3.2.2 Nicht formalisierbare Reduktionen

Nicht formalisierbare Reduktionen werden nicht angewendet.

3.3 Verwendung des Verbbregisters

3.3.1 erklärt, wie im Verbbregister nicht vorhandene Verben eindeutig einem Cluster zuzuordnen sind. 3.3.2 beschreibt die aus der computergestützten Analyse des kroatischen Verbbbestandes gewonnenen linguistischen Informationen.

3.3.1 Zuordnung eines beliebigen Verbs zu seinem Mustercluster

Die Zuordnung eines beliebigen Verbs richtet sich nach dem in II.3.3.1 vorgestellten Muster.

3.3.2 Gewinnung linguistischer Informationen aus dem Verbbregister

Die Data-Mining-Analyse liefert dem Linguisten neue, ohne Computerhilfe kaum zu gewinnende Resultate, die in Grammatiken so nicht zu finden sind.

Nachfolgend wird das Mengengerüst je Flexionstyp und Aspekt aufgeführt. Die Auflistung ist danach sortiert. Zu jeder derartigen Kombination werden höchstens drei Beispiele genannt. Dabei handelt es sich vorzugsweise um Simplicia. Nur wenn einer Kombination weniger als drei Simplicia zugeordnet sind, wer-

den präfigierte Verben aufgelistet. Sind einer Kombination drei oder weniger Verben zugeordnet, werden alle Verben dieser Kombination genannt.

Die Unterscheidung, ob ein Verb Simplex oder präfigiertes Verb ist, ist teilweise Interpretationssache (cf. II.1.2.2). Unabhängig davon sind Mengenrelationen ganz klar erkennbar.

Flexionstyp	Aspekt	Anzahl der Simplicia	Anzahl der Verben	Beispielverben zu den Kombinationen aus Flexionstyp und Aspekt
1at	ipf	312	951	igrati, birati, obećavati
1at	pf	14	213	dokončati, sputati, smrskati
1at	pf/ipf	49	53	strijeljati, večerati, aplaudirati
1 SUMME		375	1217	
2it	ipf	116	130	oblačiti, točiti, općiti
2it	pf	12	312	udariti, vjeriti, turiti
2it	pf/ipf	2	2	noćiti, tužiti
2bitmitpitvit	ipf	44	46	galamiti, lijepiti, slaviti
2bitmitpitvit	pf	15	180	snimiti, javiti, staviti
2bitmitpitvit	pf/ipf	1	1	cijepiti
2citkit	pf	1	8	baciti, od-baciti, u-baciti
2dit	ipf	23	38	hladiti, raditi, tvrditi
2dit	pf	5	88	dojaditi, iznenaditi, ozlijediti
2dit	pf/ipf	0	1	po-hoditi
2gitzit	ipf	8	33	nalaziti, gaziti, paziti
2gitzit	pf	2	21	gmiziti, suziti, u-groziti
2hitsit	ipf	10	23	gasiti, visiti, nositi
2hitsit	pf	2	25	spasiti, objesiti, po-mrsiti
2lit	ipf	21	21	paliti, šaliti, hvaliti
2lit	pf	1	51	dozvoliti, na-valiti, pre-soliti
2nit	ipf	19	21	raniti, braniti, puniti
2nit	pf	1	70	raniti, iz-mijeniti, iz-goniti
2nit	pf/ipf	1	1	cariniti
2slit	ipf	1	1	misliti
2slit	pf	0	8	za-misliti, pre-misliti, predo-misliti
2tit	ipf	15	16	mlatiti, patiti, sramotiti
2tit	pf	9	59	platiti, obratiti, vratiti,
2tit-šč	ipf	3	4	čistiti, gostiti, žalostiti
2tit-šč	pf	0	9	o-čistiti, po-gostiti, pri-čvrstiti
2tit-šč	pf/ipf	1	1	častiti
2tit-št	ipf	2	2	mastiti, koristiti

Flexionstyp	Aspekt	Anzahl der Simplicita	Anzahl der Verben	Beispielverben zu den Kombinationen aus Flexionstyp und Aspekt
2tit-št	pf	2	31	oprostiti, pustiti, pre-krstiti
2tit-št	pf/ipf	1	1	krstiti
2znit	ipf	1	1	prazniti
2znit	pf	1	3	kazniti, sa-blazniti, is-prazniti
2snit	pf	0	4	ob-jasniti, iz-jasniti, za-kasniti
2snit	pf/ipf	1	1	kasniti
2et	ipf	2	2	vreti, zreti
2et	pf	0	3	sa-zreti, pre-zreti, do-zreti
2at	ipf	24	27	klečati, trčati, zviždati
2at	pf	0	15	odležati, o-trčati, na-brojati
2at	pf/ipf	0	1	sa-držati
2 SUMME		347	1261	
3at	ipf	18	51	grebati, kašljati, derati
3at	pf	0	6	is-kašljati, o-derati, raz-derati
3et-n	pf	4	9	načeti, početi, zapeti
3catkat	ipf	20	54	natjecati, micati, vikati
3catkat	pf	0	5	za-plakati, za-nijekati, do-vikati
3dat	ipf	1	2	glodati, iz-glodati
3dat	pf	0	1	o-glodati
3gatzat	ipf	16	50	lagati, mazati, rezati
3gatzat	pf	0	34	pri-kazati, po-lagati, is-klizati
3gatzat	pf/ipf	1	1	vezati
3hatsat	ipf	14	16	jahati, kihati, pisati
3hatsat	pf	0	16	pro-duhati, is-puhati, pre-pisati
3hatsat	pf/ipf	2	2	žigosati, krunisati
3tat	ipf	16	46	metati, kretati, drhtati
3ijet	pf	0	12	na-nijeti, od-nijeti, po-nijeti
3jat	ipf	5	19	kajati, lajati, smijati
3jat	pf	0	6	po-kajati, o-stajati, i-smijati
3klet	ipf	1	1	kleti
3klet	pf	0	2	za-kleti, pro-kleti
3et-m	pf	2	7	oteti, uzeti, pre-uzeti
3at-n	pf	2	15	stati, rastati, sa-stati
3ut-p	pf	0	8	ra-suti, pre-suti, pro-suti
3rat	ipf	3	3	brati, prati, srati
3rat	pf	0	9	sa-brati, pre-brati, u-brati
3rijet	pf	3	12	doprijeti, uprijeti, prostrijeti
3ut	ipf	7	7	gladnuti, brinuti, mrznuti
3ut	pf	90	222	viknuti, trgnuti, puknuti

Flexionstyp	Aspekt	Anzahl der Simplicia	Anzahl der Verben	Beispielverben zu den Kombinationen aus Flexionstyp und Aspekt
3ut	pf/ipf	1	1	greznuti
3vat	ipf	1	1	zvati
3vat	pf	0	8	na-zvati, po-zvati, pro-zvati
3 SUMME		207	626	
4bjetmjetpjetvjet	ipf	1	1	smjeti
4bjetmjetpjetvjet	pf	0	4	sporaz-umjeti, pri-spjeti, u-spjeti
4bjetmjetpjetvjet	pf/ipf	0	1	raz-umjeti
4djet	pf	1	1	snabdjeti
4 SUMME		2	7	
5bjetmjetpjetvjet	ipf	6	6	grmjeteri, kipjeti, živjeti
5bjetmjetpjetvjet	pf	1	11	izumjeti, pre-kipjeti, pre-trpjeti
5djet	ipf	6	7	sjedjeti, štedjeti, smrdjeti
5djet	pf	0	11	o-sijedjeti, pri-štedjeti, pred-vidjeti
5djet	pf/ipf	1	1	vidjeti
5hjetsjet	ipf	1	1	visjeti
5jet	ipf	1	1	gorjeti
5jet	pf	0	4	po-modrjeti, po-gorjeti, iz-gorjeti
5ljet	ipf	4	5	željeti, boljeti, voljeti
5ljet	pf	0	5	za-željeti, o-boljeti, naj-voljeti
5njet	ipf	3	3	zelenjeti, crvenjeti, kopnjeti
5njet	pf	0	3	po-crvenjeti, po-tamnjeti, po-crnjeti
5tjet	ipf	4	4	letjeti, drhtjeti, vrtjeti
5tjet	pf	0	8	od-letjeti, po-hitjeti, pre-šutjeti
5 SUMME		27	70	
6avat	ipf	1	18	davati, za-davati, pri-znavati
6ivatovat	ipf	67	523	noćivati, kupovati, žrtvovati
6ivatovat	pf	5	17	obradovati, zakovati, darovati
6ivatovat	pf/ipf	13	16	darivati, silovati, savjetovati
6itut	ipf	5	6	biti, piti, kriti
6itut	pf	9	57	umiti, obaviti, obuti
6itut	pf/ipf	1	1	čuti
6klat	ipf	1	1	klati
6klat	pf	0	1	za-klati
6at	ipf	2	2	hramati, kapati
6ljet	ipf	1	1	mljeti
6ljet	pf	0	1	sa-mljeti
6slat	ipf	1	1	slati

Flexionstyp	Aspekt	Anzahl der Simplicia	Anzahl der Verben	Beispielverben zu den Kombinationen aus Flexionstyp und Aspekt
6slat	pf	1	4	poslati, ras-poslati, ot-poslati
6 SUMME		107	649	
7st-b	ipf	3	3	grepsti, dupsti, zepsti
7st-b	pf	0	2	po-grepsti, iz-dupsti
7st-d	ipf	3	3	krasti, jesti, presti
7st-d	pf	2	24	djesti, o-krasti, na-vesti
7st-dn	ipf	1	1	pasti
7st-dn	pf	1	12	sjesti, na-pasti, pro-pasti
7st-s	ipf	1	1	tresti
7st-s	pf	1	8	pasti, po-tresti, za-tresti
7st-t	ipf	3	3	cvasti, plesti, mesti
7st-t	pf	0	6	pro-cvasti, za-plesti, po-mesti
7st-st	ipf	1	1	rasti
7st-st	pf	0	6	na-rasti, za-rasti, iz-rasti
7st-tn	pf	1	2	sresti, su-sresti
7st-z	ipf	3	3	vesti, gristi, musti
7st-z	pf	0	11	od-vesti, za-gristi, pro-gristi
7 SUMME		20	86	
8ć-k	ipf	4	4	peći, teći, tući
8ć-k	pf	2	27	obući, svući, pri-vući
8ć-đ	pf	8	23	naći, saći, poći
8ć-g	ipf	2	2	žeći, strići
8ć-gn	ipf	2	2	leći, vrći
8ć-gn	pf	4	21	dići, stići, privići
8ć-d	ipf	1	1	ići
8ć-kn	pf	6	23	kleći, reći, pući
8 SUMME		29	103	
SUMME		1114	4019	

Nachfolgend werden morphologisch interessante homogene Cluster aufgeführt. Dabei handelt es sich um Cluster, die mehr als ein Simplex enthalten und keinem Gemeincluster zugeordnet sind.

Cluster	Flexionstyp	Anzahl Simplicia	Bespiele (Simplicia)
taći	8ć-kn	2	taći, spotaći
vući	8ć-k	2	vući, svući
tjecati	3catkat	6	natjecati, pretjecati, potjecati
ečati	2at	2	klečati, zvečati
ahati	3hatsat	2	jahati, mahati
ihati	3hatsat	2	kihati, njihati
ojati	2at	2	brojati, stojati
derati	3at	2	derati, žderati
risati	3hatsat	3	risati, brisati, mirisati
retati	3tat	2	kretati, sretati
ptati	3tat	2	šaptati, treptati
ištati	2at	3	pištati, vrištati, tištati
ovati	6ivatovat	62	trebovati, dobovati, božićovati
juvati	6ivatovat	2	kljuvati, pljuvati
azati	3gatzat	2	kazati, mazati
ezati	3gatzat	3	rezati, sezati, vezati
lizati	3gatzat	2	lizati, klizati
ežati	2at	4	bježati, ležati, režati
četi	3et-n	2	načeti, početi
edjeti	5djet	2	sjedjeti, štedjeti
idjeti	5djet	2	stidjeti, vidjeti
rijeti	3rijet	3	doprijeti, uprijeti, prostrijeti
tjeti	5tjet	4	letjeti, drhtjeti, vrtjeti
vjeti	5bjetmjjetpjetvjjet	2	živjeti, vrvjeti
eljeti	5ljet	2	bijeljeti, željeti
oljeti	5ljet	2	boljeti, voljeti
njeti	5njet	3	zelenjeti, crvenjeti, kopnjeti
peti	3et-n	2	zapeti, raspeti
reti	2et	2	vreti, zreti
abiti	2bitmitpitvit	4	slabiti, rabiti, grabiti
obiti	2bitmitpitvit	3	prisposodobiti, sukobiti, drobiti
ubiti	2bitmitpitvit	5	dubiti, gubiti, ljubiti
čiti	2it	26	prosjačiti, oblačiti, tlačiti
ćiti	2it	2	noćiti, općiti
điti	2dit	28	dojaditi, kaditi, gladiti

Cluster	Flexionstyp	Anzahl Simplicia	Bespiele (Simplicia)
jiti	2it	13	gajiti, tajiti, bojiti
aliti	2lit	7	faliti, kaliti, paliti
eliti	2lit	5	bijeliti, dijeliti, seliti
iliti	2lit	3	piliti, siliti, cviliti
iljiti	2it	3	škiljiti, piljiti, šiljiti
uljiti	2it	2	buljiti, duljiti
amiti	2bitmitpitvit	3	galamiti, mamiti, sramiti
imiti	2bitmitpitvit	3	dimiti, snimiti, primiti
jmiti	2bitmitpitvit	2	najmiti, pojmiti
omiti	2bitmitpitvit	2	lomiti, pripitomiti
zniti	2znit	2	kazniti, prazniti
apiti	2bitmitpitvit	3	ćapiti, zjapiti, vapiti
epiti	2bitmitpitvit	2	cijepiti, lijepiti
ipiti	2bitmitpitvit	2	đipiti, škripiti
rpiti	2bitmitpitvit	2	drpiti, krpiti
upiti	2bitmitpitvit	3	kupiti, lupiti, stupiti
siti	2hitsit	12	gasiti, glasiti, spasiti
ešiti	2it	5	smješiti, riješiti, griješiti
ošiti	2it	2	trošiti, pustošiti
ršiti	2it	3	kršiti, stršiti, vršiti
ušiti	2it	6	bušiti, gušiti, njušiti
atiti	2tit	9	pečatiti, latiti, mlatiti
etiti	2tit	5	primijetiti, prijetiti, sjetiti
ititi	2tit	3	hititi, kititi, štititi
otiti	2tit	3	kotiti, sramotiti, krotiti
štiti	2it	4	lašiti, pištiti, priušiti
utiti	2tit	2	slutiti, ljutiti
ziti	2gitzit	10	gaziti, paziti, kliziti
žiti	2it	18	blažiti, dražiti, tražiti
nuti	3ut	98	ganuti, kanuti, svanuti

4. Kroatisches Verbbregister

Auf den folgenden Seiten wird das durch formalisierbare Reduktionen aufbereitete Ergebnis der Analyse des kroatischen Verbbestandes vorgestellt. Um die Darstellung möglichst übersichtlich zu gestalten, erstrecken sich die Tabellenspalten über zwei gegenüberliegende Seiten.

Bezeichnung der Spalte	Bedeutung
Cluster	Bezeichnung des Clusters
Flexionstyp	Flexionstyp des Verbs (1.5.2)
Asp.	Aspekt des Verbs (1.4.2.1)
Infinitiv	Schlüsselformen der Beispielverben (1.5.1)
Präsens 3. Pl.	
Präteritum Sg. feminin	
Aorist 3. Sg.	
Imperfekt 1. Sg.	
Imperativ Sg.	
Adverbialpartizip Präteritum	
Kommentare	

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
	8ć-kn	pf	od-maći	od-maknu	od-makla
	8ć-đ	pf	naći	nađu	našla
	8ć-đ	pf	saći	sađu	sašla
homC	8ć-kn		~taći	~taknu	~takla
	8ć-đ	pf	zaći	zađu	zašla
	8ć-gn	pf	pre-bjeći	pre-bjegnu	pre-bjegla
	8ć-k	pf	ra-sjeći	ra-sjeku	ra-sjekla
	8ć-gn	ipf	leći	legnu	legla
	8ć-kn	pf	kleći	kleknu	klekla
	8ć-k	ipf	peći	peku	pekla
	8ć-kn	pf	reći	reknu	rekla
	8ć-gn	pf	pri-seći	pri-segnu	pri-segla
	8ć-k	ipf	teći	teku	tekla
	8ć-kn	pf	steći	steknu	stekla
	8ć-g	ipf	žeći	žegu	žegla
	8ć-d	ipf	ići	idu	išla
	8ć-đ	pf	na-ići	na-iđu	na-išla
	8ć-đ	pf	ob-ići	ob-iđu	ob-išla
	8ć-đ	pf	zaob-ići	zaob-iđu	zaob-išla
	8ć-đ	pf	mimo-ići	mimo-iđu	mimo-išla
	8ć-đ	pf	s-ići	s-iđu	s-išla
	8ć-đ	pf	ot-ići	ot-iđu	ot-išla
	8ć-đ	pf	raz-ići	raz-iđu	raz-išla
	8ć-đ	pf	iz-ići	iz-iđu	iz-išla
	8ć-đ	pf	proiz-ići	proiz-iđu	proiz-išla
	8ć-gn	pf	dići	dignu	digla
	8ć-kn	pf	iz-nići	iz-niknu	iz-nikla
	8ć-đ	pf	prići	priđu	prišla
	8ć-g	ipf	strići	strigu	strigla
	8ć-gn	pf	stići	stignu	stigla
	8ć-kn	pf	na-vići	na-viknu	na-vikla
	8ć-gn	pf	privići	privignu	privigla
	8ć-đ	pf	doći	dođu	došla
	8ć-gn	pf	pomoći	pomognu	pomogla
	8ć-đ	pf	poći	pođu	pošla
	8ć-đ	pf	proći	prođu	prošla
	8ć-gn	ipf	vrći	vrgnu	vrgla
	8ć-đ	pf	ući	uđu	ušla
	8ć-k	pf	obući	obuku	obukla
	8ć-kn	pf	pući	puknu	pukla
	8ć-k	ipf	tući	tuku	tukla
homC	8ć-k		~vući	~vuku	~vukla
basC	1at		~ati	~aju	~ala
	3at	ipf	grebati	grebu	grebala
homC	3catkat		~tjecati	~tječu	~tjecala

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
od-mače		od-maci	od-makavši	
nađe		nađi	našavši	
sađe		sađi	sašavši	
~tače	~tačah	~taci	~takavši	taći, spotaći
zađe		zađi	zašavši	
pre-bježe		pre-bjezi	pre-bjegavši	
ra-sječe		ra-sjeci	ra-sjekavši	
leže	ležah	lezi	legavši	
kleče		kleci	klekavši	
peče	pečah	peci	pekavši	
reče		reci	rekavši	
pri-seže		pri-sezi	pri-segavši	
teče	tečah	teci	tekavši	
steče		steci	stekavši	
žeže	žežah	žezi	žegavši	
ide	idah	idi	išavši	
na-iđe		na-iđi	na-išavši	
ob-iđe		ob-iđi	ob-išavši	
zaob-iđe		zaob-iđi	zaob-išavši	
mimo-iđe		mimo-iđi	mimo-išavši	
s-iđe		s-iđi	s-išavši	
ot-iđe		ot-iđi	ot-išavši	
raz-iđe		raz-iđi	raz-išavši	
iz-iđe		iz-iđi	iz-išavši	
proiz-iđe		proiz-iđi	proiz-išavši	
dize		dizi	digavši	
iz-niče		iz-nici	iz-nikavši	
piđe		piđi	prišavši	
striže	strižah	strizi	strigavši	
stiže		stizi	stigavši	
na-viče		na-vici	na-vikavši	
priviže		privizi	privigavši	
dođe		dođi	došavši	
pomože		pomozi	pomogavši	
pođe		pođi	pošavši	
prođe		prođi	prošavši	
vrže	vržah	vrzi	vrgavši	
uđe		uđi	ušavši	
obuče		obuci	obukavši	
puče		puci	pukavši	
tuče	tučah	tuci	tukavši	
~vuče	~vučah	~vuci	~vukavši	vući, svući
~a	~ah	~aj	~avši	~ati \ ~ivati
greba	grebah	grebi	grebavši	
~tjeca	~tjecah	~tječi	~tjecavši	natjecati, pretjecati

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
x	1at	ipf	pečati	pečaju	pečala
x	1at	ipf	klicati	klicaju	klicala
	3catkat	ipf	micati	miču	micala
	3catkat	ipf	nicati	niču	nicala
	3catkat	ipf	od-ricati	od-riču	od-ricala
	3catkat	ipf	ticati	tiču	ticala
homC	2at		~ečati	~eče	~ečala
	2at	ipf	cičati	ciče	cičala
x	1at	ipf	veličati	veličaju	veličala
	2at	ipf	kričati	kriče	kričala
	2at	ipf	krčati	krče	krčala
x	1at	pf	o-gorčati	o-gorčaju	o-gorčala
	2at	ipf	trčati	trče	trčala
x	1at	ipf	bučati	bučaju	bučala
	2at	ipf	čučati	čuče	čučala
	2at	ipf	zvučati	zvuče	zvučala
	3dat	ipf	glodati	glođu	glodala
	2at	ipf	zviždati	zvižde	zviždala
	3gatzat	ipf	lagati	lažu	lagala
	3gatzat	ipf	pomagati	pomažu	pomagala
x	1at	ipf	tragati	tragaju	tragala
	3gatzat	ipf	vagati	važu	vagala
	3gatzat	ipf	s-lijegati	s-liježu	s-lijegala
	3gatzat	ipf	strugati	stružu	strugala
homC	3hatsat		~ahati	~ašu	~ahala
homC	3hatsat		~ihati	~išu	~ihala
	3hatsat	ipf	duhhati	dušu	duhala
x	1at	ipf	kuhati	kuhaju	kuhala
	3hatsat	ipf	puhati	pušu	puhala
x	1at	ipf	od-gajati	od-gajaju	od-gajala
	3jat	ipf	kajati	kaju	kajala
	3jat	ipf	lajati	laju	lajala
	3jat	ipf	na-stajati	na-staju	na-stajala
x	1at	ipf	na-bijati	na-bijaju	na-bijala
	3jat	ipf	smijati	smiju	smijala
	3jat	ipf	brijati	briju	brijala
	3jat	ipf	grijati	griju	grijala
x	1at	ipf	prijati	prijaju	prijala
homC	2at		~ojati	~oje	~ojala
x	1at	ipf	bujati	bujaju	bujala
	2at	ipf	hujati	huje	hujala
	2at	ipf	brujati	bruje	brujala
	2at	ipf	zujati	zuje	zujala
x	1at	ipf	ljubakati	ljubakaju	ljubakala
	3catkat	ipf	skakati	skaču	skakala

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
peca	pecah	pecaj	pecavši	
klica	klicah	klicaj	klicavši	
mica	micah	miči	micavši	
nica	nicah	niči	nicavši	
od-rica	od-ricah	od-riči	od-ricavši	
tica	ticah	tiči	ticavši	
~eča	~ečah	~eči	~ečavši	klečati, zvečati
ciča	cičah	ciči	cičavši	
veliča	veličah	veličaj	veličavši	
kriča	kričah	kriči	kričavši	
krča	krčah	krči	krčavši	
o-gorča		o-gorčaj	o-gorčavši	
trča	trčah	trči	trčavši	
buča	bučah	bučaj	bučavši	
čuča	čučah	čuči	čučavši	
zvuča	zvučah	zvuči	zvučavši	
gloda	glodah	glodī	glodavši	
zvižda	zviždah	zviždi	zviždavši	
laga	lagah	laži	lagavši	
pomaga	pomagah	pomaži	pomagavši	
traga	tragah	tragaj	tragavši	
vaga	vagah	važi	vagavši	
s-lijega	s-lijegah	s-liježi	s-lijegavši	
struga	strugah	struži	strugavši	
~aha	~ahah	~aši	~ahavši	jahati, mahati
~iha	~ihah	~iši	~ihavši	kihati, njihati
duha	duhah	duši	duhavši	
kuha	kuhah	kuhaj	kuhavši	
puha	puhah	puši	puhavši	
od-gaja	od-gajah	od-gajaj	od-gajavši	
kaja	kajah	kaj	kajavši	
laja	lajah	laj	lajavši	
na-staja	na-stajah	na-staj	na-stajavši	
na-bija	na-bijah	na-bijaj	na-bijavši	
smija	smijah	smij	smijavši	
brija	brijah	brij	brijavši	
grija	grijah	grij	grijavši	
prija	prijah	prijaj	prijavši	
~oja	~ojah	~oji	~ojavši	brojati, stojati
buja	bujah	bujaj	bujavši	
huja	hujah	huji	hujavši	
bruja	brujah	bruji	brujavši	
zuja	zujah	zuji	zujavši	
ljubaka	ljubakah	ljubakaj	ljubakavši	
skaka	skakah	skači	skakavši	

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
	3catkat	ipf	plakati	plaču	plakala
	3catkat	ipf	na-makati	na-maču	na-makala
	3catkat	ipf	na-takati	na-taču	na-takala
	3catkat	ipf	žvakati	žvaču	žvakala
	3catkat	ipf	nijekati	niječu	nijekala
x	1at	ipf	slikati	slikaju	slikala
	3catkat	ipf	rikati	riču	rikala
	3catkat	ipf	tikati	tiču	tikala
	3catkat	ipf	vikati	viču	vikala
	3catkat	ipf	lokati	loču	lokala
	3catkat	ipf	srkati	srču	srkala
	3catkat	ipf	jaukati	jauču	jaukala
x	1at	pf	o-žbukati	o-žbukaju	o-žbukala
	3catkat	ipf	gukati	guču	gukala
	3catkat	pf	na-sukati	na-suču	na-sukala
	6klat	ipf	klati	kolju	klala
	6slat	ipf	slati	šalju	slala
x	1at	ipf	veslati	veslaju	veslala
	6slat	pf	poslati	pošalju	poslala
x	1at	ipf	valjati	valjaju	valjala
	3at	ipf	raza-šiljati	raza-šilju	raza-šiljala
x	1at	ipf	iza-šiljati	iza-šiljaju	iza-šiljala
	3at	ipf	po-šiljati	po-šilju	po-šiljala
	3at	ipf	kašljati	kašlju	kašljala
	6at	ipf	hramati	hramlju	hramala
x	1at	ipf	pro-ganjati	pro-ganjaju	pro-ganjala
x	1at	ipf	mijenjati	mijenjaju	mijenjala
	3at	ipf	penjati	penju	penjala
	3at	ipf	stenjati	stenju	stenjala
homC	3at		~injati	~inju	~injala
	6at	ipf	kapati	kaplju	kapala
	1at	ipf	is-kapati	is-kapaju	is-kapala
	3rat	ipf	brati	beru	brala
homC	3at		~derati	~deru	~derala
x	1at	ipf	birati	biraju	birala
	3at	ipf	oda-birati	oda-biru	oda-birala
	3at	ipf	sa-birati	sa-biru	sa-birala
	3at	ipf	raza-birati	raza-biru	raza-birala
	3at	ipf	iza-birati	iza-biru	iza-birala
	3at	ipf	po-birati	po-biru	po-birala

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
plaka	plakah	plači	plakavši	
na-maka	na-makah	na-mači	na-makavši	
na-taka	na-takah	na-tači	na-takavši	
žvaka	žvakah	žvači	žvakavši	
nijeka	nijekah	niječi	nijekavši	
slika	slikah	slikaj	slikavši	
rika	rikah	riči	rikavši	
tika	tikah	tiči	tikavši	
vika	vikah	viči	vikavši	
loka	lokah	loči	lokavši	
srka	srkah	srči	srkavši	
jauka	jaukah	jauči	jaukavši	
o-žbuka		o-žbukaj	o-žbukavši	
guka	gukah	guči	gukavši	
na-suka		na-suči	na-sukavši	
kla	klah	kolji	klavši	
sla	slah	šalji	slavši	
vesla	veslah	veslaj	veslavši	
posla		pošalji	poslavši	
valja	valjah	valjaj	valjavši	
raza-šilja	raza-šiljah	raza-šilji	raza-šiljavši	
iza-šilja	iza-šiljah	iza-šiljaj	iza-šiljavši	
po-šilja	po-šiljah	po-šilji	po-šiljavši	
kašlja	kašljah	kašlji	kašljavši	
hrama	hramah	hramlji	hramavši	
pro-ganja	pro-ganjah	pro-ganjaj	pro-ganjavši	
mijenja	mijenjah	mijenjaj	mijenjavši	
penja	penjah	penji	penjavši	
stenja	stenjah	stenji	stenjavši	
~inja		~inji	~injavši	na-činjati, na-ginjati, za-klinjati, o-pominjati, za-pinjati, o-dapinjati, ošinjati
	~injah			
kapa	kapah	kaplji	kapavši	
is-kapa	is-kapah	is-kapaj	is-kapavši	
bra	brah	beri	bravši	
~dera	~derah	~deri	~deravši	derati, žderati
bira	birah	biraj	biravši	
oda-bira	oda-birah	oda-biri	oda-biravši	
sa-bira	sa-birah	sa-biri	sa-biravši	
raza-bira	raza-birah	raza-biri	raza-biravši	
iza-bira	iza-birah	iza-biri	iza-biravši	
po-bira	po-birah	po-biri	po-biravši	

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
	3at	ipf	u-birati	u-biru	u-birala
	3at	ipf	pre-birati	pre-biru	pre-birala
	3at	ipf	pro-dirati	pro-diru	pro-dirala
	3at	ipf	obamirati	obamiru	obamirala
x	1at	pf/ipf	animirati	animiraju	animirala
	3at	ipf	umirati	umiru	umirala
	3at	ipf	prepirati	prepiru	prepirala
x	1at	ipf	dopirati	dopiraju	dopirala
	3at	ipf	upirati	upiru	upirala
x	1at	pf/ipf	eksploatirati	eksploatiraju	eksploatirala
	3at	ipf	otirati	otiru	otirala
	3at	ipf	zastirati	zastiru	zastirala
	1at	pf/ipf	protestirati	protestiraju	protestirala
	3at	ipf	prostirati	prostiru	prostirala
	3at	ipf	utirati	utiru	utirala
	3at	ipf	oba-zirati	oba-ziru	oba-zirala
	3at	ipf	orati	oru	orala
	3rat	ipf	prati	peru	prala
	3rat	ipf	srati	seru	srala
x	1at	ipf	talasati	talasaju	talasala
	3hatsat	ipf	pasati	pašu	pasala
	3hatsat	ipf	plesati	plešu	plesala
x	1at	ipf	za-tresati	za-tresaju	za-tresala
	3hatsat	ipf	tesati	tešu	tesala
	3hatsat	ipf	disati	dišu	disala
	3hatsat	pf/ipf	krunisati	krunišu	krunisala
	3hatsat	ipf	pisati	pišu	pisala
homC	3hatsat		~risati	~rišu	~risala
x	1at	ipf	sisati	sisaju	sisala
	3hatsat	pf	i-sisati	i-sišu	i-sisala
	3hatsat	pf	u-sisati	u-sišu	u-sisala
	3hatsat	pf/ipf	žigosati	žigošu	žigosala
	3at	ipf	češati	češu	češala
	2at	ipf	stršati	strše	stršala
	3tat	ipf	na-lijetati	na-lijeću	na-lijetala
x	1at	ipf	cvjetati	cvjetaju	cvjetala
	3tat	ipf	zveketati	zvekeću	zveketala
	3tat	ipf	za-pletati	za-pleću	za-pletala
	3tat	ipf	metati	meću	metala
	3tat	ipf	trepetati	trepeću	trepetala
homC	3tat		~retati	~reću	~retala
	3tat	ipf	drhtati	drhću	drhtala
	3tat	ipf	uplitati	upliču	uplitala
	3tat	ipf	zvektati	zvekću	zvektala
	3tat	ipf	roktati	rokću	roktala

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
u-bira	u-birah	u-biri	u-biravši	
pre-bira	pre-birah	pre-biri	pre-biravši	
pro-dira	pro-dirah	pro-điri	pro-diravši	
obamira	obamirah	obamiri	obamiravši	
animira	animirah	animiraj	animiravši	
umira	umirah	umiri	umiravši	
prepira	prepirah	prepiri	prepiravši	
dopira	dopirah	dopiraj	dopiravši	
upira	upirah	upiri	upiravši	
eksploatira	eksploatirah	eksploatiraj	eksploatiravši	
otira	otirah	otiri	otiravši	
zastira	zastirah	zastiri	zastiravši	
protestira	protestirah	protestiraj	protestiravši	
prostira	prostirah	prostiri	prostiravši	
utira	utirah	utiri	utiravši	
oba-zira	oba-zirah	oba-ziri	oba-ziravši	
ora	orah	ori	oravši	
pra	prah	peri	pravši	
sra	srah	seri	sravši	
talasa	talasah	talasaj	talasavši	
pasa	pasah	paši	pasavši	
plesa	plesah	pleši	plesavši	
za-tresa	za-tresah	za-tresaj	za-tresavši	
tesa	tesah	teši	tesavši	
disa	disah	diši	disavši	
krunisa	krunisah	kruniši	krunisavši	
pisa	pisah	piši	pisavši	
~risa	~risah	~riši	~risavši	risati, brisati, mirisati
sis	sisah	sisaj	sisavši	
i-sis		i-siši	i-sisavši	
u-sis		u-siši	u-sisavši	
žigosa	žigosah	žigoši	žigosavši	
češa	češah	češi	češavši	
strša	stršah	strši	stršavši	
na-lijeta	na-lijetah	na-lijeći	na-lijetavši	
cvjeta	cvjetah	cvjetaj	cvjetavši	
zveketa	zveketah	zvekeći	zveketavši	
za-pleta	za-pletah	za-pleći	za-pletavši	
meta	metah	meći	metavši	
trepeta	trepetah	trepeći	trepetavši	
~reta	~retah	~reći	~retavši	kretati, sretati
drhta	drhtah	drhći	drhtavši	
uplita	uplitah	uplići	uplitavši	
zvekta	zvektah	zvekći	zvektavši	
rokta	roktah	rokći	roktavši	

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
	2at	ipf	buktati	bukte	buktala
	3tat	ipf	grgotati	grgoću	grgotala
homC	3tat		~ptati	~pću	~ptala
	3tat	ipf	obrtati	obrću	obrtala
x	1at	ipf	crtati	crtaju	crtala
	3tat	ipf	o-grtati	o-grću	o-grtala
	3tat	ipf	po-srtati	po-srću	po-srtala
	3tat	ipf	za-vrtati	za-vrću	za-vrtala
	3at-n	pf	stati	stanu	stala
	3at-n	pf	rastati	rastanu	rastala
x	1at	ipf	na-rastati	na-rastaju	na-rastala
x	1at	ipf	praštati	praštaju	praštala
homC	2at		~ištati	~ište	~ištala
	2at	ipf	pljuštati	pljušte	pljuštala
x	1at	ipf	gutati	gutaju	gutala
	3tat	ipf	škrgutati	škrguću	škrgutala
	3tat	ipf	skakutati	skakuću	skakutala
	3tat	ipf	šaputati	šapuću	šaputala
x	1at	ipf	is-probavati	is-probavaju	is-probavala
	6avat	ipf	davati	daju	davala
x	1at	ipf	sa-vladavati	sa-vladavaju	sa-vladavala
	6avat	ipf	pre-udavati	pre-udaju	pre-udavala
	6avat	ipf	raza-znavati	raza-znaju	raza-znavala
basC2	6ivatovat		~ivati	~uju	~ivala
	1at	ipf	počivati	počivaju	počivala
x	6ivatovat	ipf	o-plakivati	o-plakuju	o-plakivala
	1at	ipf	zakivati	zakivaju	zakivala
	1at	ipf	okivati	okivaju	okivala
	1at	ipf	potkivati	potkivaju	potkivala
	1at	ipf	s-livati	s-livaju	s-livala
x	6ivatovat	ipf	is-cjelivati	is-cjeluju	is-cjelivala
	1at	ipf	plivati	plivaju	plivala
	1at	ipf	u-mivati	u-mivaju	u-mivala
	1at	ipf	za-snivati	za-snivaju	za-snivala
	1at	ipf	sa-krivati	sa-krivaju	sa-krivala
	1at	ipf	šivati	šivaju	šivala
	1at	ipf	zaptivati	zaptivaju	zaptivala
	1at	ipf	na-zivati	na-zivaju	na-zivala
x	6ivatovat	ipf	kazivati	kazuju	kazivala
	1at	ipf	zazivati	zazivaju	zazivala
	1at	ipf	o-porezivati	o-porezivaju	o-porezivala
	1at	ipf	uživati	uživaju	uživala
homC	6ivatovat		~ovati	~uju	~ovala
	3at	ipf	rvati	rvu	rvala

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
bukta	buktah	bukti	buktavši	
grgota	grgotah	grgoći	grgotavši	
~pta	~ptah	~pći	~ptavši	šaptati, treptati
obrta	obrtah	obrći	obrtavši	
crtā	crtah	crtaj	crtavši	
o-grta	o-grtah	o-grći	o-grtavši	
po-srta	po-srtah	po-srći	po-srtavši	
za-vrta	za-vrtah	za-vrći	za-vrtavši	
sta		stani	stavši	
rasta		rastani	rastavši	
na-rasta	na-rastah	na-rastaj	na-rastavši	
prašta	praštah	praštaj	praštavši	
~išta	~ištah	~išti	~ištavši	pištati, vrištati, tištati
pljušta	pljuštah	pljušti	pljuštavši	
guta	gutah	gutaj	gutavši	
škrgota	škrgotah	škrgući	škrgotavši	
skakuta	skakutah	skakući	skakutavši	
šaputa	šaputah	šapući	šaputavši	
is-probava	is-probavah	is-probavaj	is-probavavši	
dava	davah	davaj	davavši	
sa-vladava	sa-vladavah	sa-vladavaj	sa-vladavavši	
pre-udava	pre-udavah	pre-udavaj	pre-udavavši	
raza-znava	raza-znavah	raza-znavaj	raza-znavavši	
~iva	~ivah	~uj	~ivavši	~ivati
počiva	počivah	počivaj	počivavši	
o-plakiva	o-plakivah	o-plakuj	o-plakivavši	
zakiva	zakivah	zakivaj	zakivavši	
okiva	okivah	okivaj	okivavši	
potkiva	potkivah	potkivaj	potkivavši	
s-liva	s-livah	s-livaj	s-livavši	
is-cjeliva	is-cjelivah	is-cjeluj	is-cjelivavši	
pliva	plivah	plivaj	plivavši	
u-miva	u-mivah	u-mivaj	u-mivavši	
za-sniva	za-snivah	za-snivaj	za-snivavši	
sa-kriva	sa-krivah	sa-krivaj	sa-krivavši	
šiva	šivah	šivaj	šivavši	
zaptiva	zaptivah	zaptivaj	zaptivavši	
na-ziva	na-zivah	na-zivaj	na-zivavši	
kaziva	kazivah	kazuj	kazivavši	
zaziva	zazivah	zazivaj	zazivavši	
o-poreziva	o-porezivah	o-porezivaj	o-porezivavši	
uživa	uživah	uživaj	uživavši	
~ova	~ovah	~uj	~ovavši	viele
rva	rvah	rvi	rvavši	

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
homC	6ivatovat		~ljuvati	~ljuju	~ljuvala
	3vat	ipf	zvati	zovu	zvala
homC	3gatzat		~azati	~ažu	~azala
homC	3gatzat		~ezati	~ežu	~ezala
	3gatzat	ipf	dizati	dižu	dizala
homC	3gatzat		~lizati	~ližu	~lizala
	3gatzat	ipf	gmizati	gmižu	gmizala
	3gatzat	ipf	nizati	nižu	nizala
x	1at	ipf	na-grizati	na-grizaju	na-grizala
	3gatzat	ipf	stizati	stižu	stizala
	3gatzat	ipf	rzati	ržu	rzala
	3gatzat	ipf	puzati	pužu	puzala
x	1at	ipf	opažati	opažaju	opažala
homC	2at		~ežati	~eže	~ežala
	2at	ipf	držati	drže	držala
homC	3et-n		~četi	~čnu	~čela
	5bjetmjetpjetvjet	ipf	svrbjeti	svrbe	svrbjela
	5djet	ipf	bdjeti	bde	bdjela
	4djet	pf	snabdjeti	snabdiju	snabdjela
	5djet	ipf	lebdjeti	lebde	lebdjela
homC	5djet		~edjeti	~ede	~edjela
homC	5djet		~idjeti	~ide	~idjela
	5djet	ipf	smrdjeti	smrde	smrdjela
	3ijet	pf	na-nijeti	na-nesu	na-nijela
homC	3rijet		~rijeti	~ru	~rla
	5bjetmjetpjetvjet	ipf	grmjeti	grme	grmjela
	4bjetmjetpjetvjet	ipf	smjeti	smiju	smjela
	4bjetmjetpjetvjet	pf/ipf	raz-umjeti	raz-umiju	raz-umjela
	5bjetmjetpjetvjet	pf	izumjeti	izume	izumjela
	5bjetmjetpjetvjet	pf	is-hlapjeti	is-hlape	is-hlapjela
	5bjetmjetpjetvjet	pf	o-slijepjeti	o-slijepe	o-slijepjela
	5bjetmjetpjetvjet	ipf	kipjeti	kipe	kipjela
	5bjetmjetpjetvjet	ipf	trpjeti	trpe	trpjela
	4bjetmjetpjetvjet	pf	pri-spjeti	pri-spiju	pri-spjela

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
~ljuva	~ljuvah	~ljuj	~ljuvavši	kljuvati, pljuvati
zva	zvah	zovi	zvavši	
~aza	~azah	~aži	~azavši	kazati, mazati
~eza		~eži	~ezavši	rezati, u-prezati, sezati, na-tezati, vezati
	~ezah			
diza	dizah	diži	dizavši	
~liza	~lizah	~liži	~lizavši	lizati, klizati
gmiza	gmizah	gmiži	gmizavši	
niza	nizah	niži	nizavši	
na-griza	na-grizah	na-grizaj	na-grizavši	
stiza	stizah	stiži	stizavši	
rza	rzah	rži	rzavši	
puza	puzah	puži	puzavši	
opaža	opažah	opažaj	opažavši	
~eža		~eži	~ežavši	bježati, ležati, režati
	~ežah			
drža	držah	drži	državši	
~če		~čni	~čevši	načeti, početi
svrbje	svrbljah	svrbi	svrbjevši	
bdje	bđah	bdi	bdjevši	
snabdje		snabdij	snabdjevši	
lebdje	lebđah	lebdi	lebdjevši	
~edje		~edi	~edjevši	o-sijedjeti, sjedjeti, štedjeti
	~eđah			
~idje	~iđah	~idi	~idjevši	stidjeti, vidjeti
smrdje	smrđah	smrdi	smrdjevši	
na-nije		na-nesi	na-nijevši	
~rije		~ri	~rijevši	pro-drijeti, oba-mrijeti, doprijeti, uprijeti, za-strijeti, prostrijeti
	~rah			
grmje	grmljah	grmi	grmjevši	
smje	smijah	smij	smjevši	
raz-umje	raz-umijah	raz-umij	raz-umjevši	
izumje		izumi	izumjevši	
is-hlapje		is-hlapi	is-hlapjevši	
o-slijepje		o-slijepi	o-slijepjevši	
kipje	kipljah	kipi	kipjevši	
trpje	trpljah	trpi	trpjevši	
pri-spje		pri-spjij	pri-spjevši	

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
homC	5jet		~rjeti	~re	~rjela
	5hjetsjet	ipf	visjeti	vise	visjela
homC	5tjet		~tjeti	~te	~tjela
homC	5bjetmjjetpjjetvjet		~vjeti	~ve	~vjela
	3klet	ipf	kleti	kunu	klela
homC	5ljet		~eljeti	~ele	~eljela
	6ljet	ipf	mljeti	melju	mljela
homC	5ljjet		~oljeti	~ole	~oljela
homC	5njet		~njeti	~ne	~njela
homC	3et-n		~peti	~pnu	~pela
homC	2et		~reti	~re	~rela
	3et-m	pf	oteti	otmu	otela
	3et-m	pf	uzeti	uzmu	uzela
	6itut	ipf	biti	biju	bila
homC	2bitmitpitvit		~abiti	~abe	~abila
	2bitmitpitvit	pf	upo-trijebiti	upo-trijebe	upo-trijebila
homC	2bitmitpitvit		~obiti	~obe	~obila
homC	2bitmitpitvit		~rbiti	~rbe	~rbila
homC	2bitmitpitvit		~ubiti	~ube	~ubila
	2citkit	pf	baciti	bace	bacila
homC	2it		~čiti	~če	~čila
homC	2it		~ćiti	~će	~ćila
homC	2dit		~diti	~de	~dila

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
~rje	~rah	~ri	~rjevši	po-modrjeti, gorjeti
visje	višah	visi	visjevši	
~tje	~ćah	~ti	~tjevši	letjeti, drhtjeti, po-hitjeti, vrtjeti, ćutjeti, pre-šutjeti
~vje	~vljah	~vi	~vjevši	živjeti, vrvjeti
kle		kuni	klevši	
~elje	~eljah	~eli	~eljevši	bijeljjeti, željeti
mlje	mljeh	melji	mljevši	
~olje	~oljah	~oli	~oljevši	boljeti, o-doljeti, voljeti
~nje	~njah	~ni	~njevši	zelenjeti, crvenjeti, po-tamnjeti, kopnjeti, po-crnjeti
~pe		~pni	~pevši	oda-peti, zapeti, raspeti
~re		~ri	~revši	vreti, zreti
ote		otmi	otevši	
uze		uzmi	uzevši	
bi	bijah	bij	bivši	
~abi	~abljah	~abi	~abivši	slabiti, rabiti, grabiti, vabiti
upo-trijebe		upo-trijebe	upo-trijebevši	
~obi	~obljah	~obi	~obivši	prisposodobiti, sukobiti, za-robiti, drobiti, one-sposobiti
~rbi	~rbljah	~rbi	~rbivši	po-grbiti, skrbiti
~ubi	~ubljah	~ubi	~ubivši	dubiti, gubiti, u-lubiti, ljubiti, od- rubiti, obrubiti, trubiti
baci		baci	bacivši	
~či	~ćah	~či	~čivši	viele
~ći	~ćah	~ći	~ćivši	o-sumnjičiti, opuno-močiti, nočiti, opčiti, one- mogućiti
~di	~đah	~di	~divši	viele

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
homC	2it		~điti	~đe	~đila
	2it	pf	od-šarafiti	od-šarafe	od-šarafila
homC	2it		~jiti	~je	~jila
	6itut	ipf	na-liti	na-liju	na-lila
homC	2lit		~aliti	~ale	~alila
	2lit	pf	za-obliti	za-oble	za-oblila
homC	2lit		~eliti	~ele	~elila
	2lit	ipf	tegliti	tegle	teglila
homC	2lit		~iliti	~ile	~ilila
	6itut	pf	doliti	doliju	dolila
	2lit	pf	raz-goliti	raz-gole	raz-golila
	2lit	pf	op-koliti	op-kole	op-kolila
	2lit	ipf	moliti	mole	molila
	6itut	pf	politi	poliju	polila
	6itut	pf	proliti	proliju	prolila
	2lit	ipf	soliti	sole	solila
	2lit	pf	dozvoliti	dozvole	dozvolila
homC	2lit		~rliti	~rle	~rlila
	2slit	ipf	misлити	misle	misлila
	2it	pf	za-posliti	za-posle	za-poslila
	2lit	ipf	svijetliti	svijetle	svijetlila
homC	2lit		~uliti	~ule	~ulila
	2it	pf	u-daljiti	u-dalje	u-daljila
homC	2it		~eljiti	~elje	~eljila
homC	2it		~iljiti	~ilje	~iljila
	2it	pf	u-zemljiti	u-zemlje	u-zemljila
	2it	ipf	grgoljiti	grgolje	grgoljila
	2lit	pf	is-poljiti	is-polje	is-poljila
	2it	pf	za-dovoljiti	za-dovolje	za-dovoljila
homC	2it		~uljiti	~ulje	~uljila
homC	2bitmitpitvit		~amiti	~ame	~amila
homC	2bitmitpitvit		~emiti	~eme	~emila
	2bitmitpitvit	ipf	jagmiti	jagme	jagmila

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
~đi	~đah	~đi	~đivši	o-međiti, o-tuđiti
od-šarafi		od-šarafi	od-šarafivši	
~ji	~jah	~ji	~jivši	viele
na-li	na-lijah	na-lij	na-livši	
~ali	~aljah	~ali	~alivši	faliti, kaliti, paliti, šaliti, taliti, na- valiti, hvaliti, žaliti
za-obli		za-obli	za-oblivši	
~eli	~eljah	~eli	~elivši	bijeliti, za-cijeliti, dijeliti, u-strijeliti, seliti, veseliti, kiseliti
tegli	tegljah	tegli	teglivši	
~ili	~iljah	~ili	~ilivši	piliti, nat-kriliti, siliti, cviliti
doli		dolij	dolivši	
raz-goli		raz-goli	raz-golivši	
op-koli		op-koli	op-kolivši	
moli	moljah	moli	molivši	
poli		polij	polivši	
proli		prolij	prolivši	
solli	soljah	solli	solivši	
dozvoli		dozvoli	dozvolivši	
~rli	~rljah	~rli	~rlivši	grliti, po-hrliti
misli	mišljah	misli	mislivši	
za-posli		za-posli	za-poslivši	
svijetli	svijetljah	svijetli	svijetlivši	
~uli	~uljah	~uli	~ulivši	guliti, is-truliti
u-dalji		u-dalji	u-daljivši	
~elji	~eljah	~elji	~eljivši	temeljiti, s-prijateljiti
~ilji	~iljah	~ilji	~iljivši	škiljiti, piljiti, šiljiti
u-zemlji		u-zemlji	u-zemljivši	
grgolji	grgoljah	grgolji	grgoljivši	
is-polji		is-polji	is-poljivši	
za-dovolji		za-dovolji	za-dovoljivši	
~ulji	~uljah	~ulji	~uljivši	na-uljiti, buljiti, duljiti
~ami	~amljah	~ami	~amivši	galamiti, mamiti, sramiti
~emi	~emljah	~emi	~emivši	lemiti, za-premiti
jagmi	jagmljah	jagmi	jagmivši	

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
homC	2bitmitpitvit		~imiti	~ime	~imila
homC	2bitmitpitvit		~jmiti	~jme	~jmila
homC	2bitmitpitvit		~omiti	~ome	~omila
	2bitmitpitvit	pf	pod-jarmiti	pod-jarme	pod-jarmila
	6itut	pf	umiti	umiju	umila
	2bitmitpitvit	ipf	glumiti	glume	glumila
	2bitmitpitvit	ipf	šumiti	šume	šumila
basC	2nit		~niti	~ne	~nila
homC	2snit		~sniti	~sne	~snila
homC	2znit		~zniti	~zne	~znila
homC	2it		~njiti	~nje	~njila
	6itut	ipf	piti	piju	pila
homC	2bitmitpitvit		~apiti	~ape	~apila
homC	2bitmitpitvit		~epiti	~epe	~epila
homC	2bitmitpitvit		~ipiti	~ipe	~ipila
	6itut	pf	opiti	opiju	opila
	2bitmitpitvit	ipf	škopiti	škope	škopila
	2bitmitpitvit	pf	pre-klopiti	pre-klope	pre-klopila
	2bitmitpitvit	ipf	škropiti	škrope	škropila
	2bitmitpitvit	ipf	topiti	tope	topila
homC	2bitmitpitvit		~rpiti	~rpe	~rpila
homC	2bitmitpitvit		~upiti	~upe	~upila
	6itut	pf	iz-riti	iz-riju	iz-rila
basC	2it		~riti	~re	~rila
	6itut	ipf	kriti	kriju	krila
homC	2hitsit		~siti	~se	~sila
	6itut	ipf	šiti	šiju	šila
	2it	pf	iš-čašiti	iš-čaše	iš-čašila

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
~imi	~imljah	~imi	~imivši	dimiti, snimiti, primiti
~jmi	~jmljah	~jmi	~jmivši	najmiti, po-zajmiti, pojmiti
~omi	~omljah	~omi	~omivši	lomiti, pripitomiti
pod-jarmi umi		pod-jarmi umij	pod-jarmivši umivši	
glumi	glumljah	glumi	glumivši	
šumi	šumljah	šumi	šumivši	
~ni	~njah	~ni	~nivši	~niti \ ~sniti \ ~zniti
~sni	~šnjah	~sni	~snivši	ob-jasniti, kasniti
~zni	~žnjah	~zni	~znivši	kazniti, sa-blazniti, prazniti
~nji	~njah	~nji	~njivši	s-manjiti, za-panjiti, tanjiti, krnjiti
pi	pijah	pij	pivši	
~api	~apljah	~api	~apivši	ćapiti, zjapiti, is-kapiti, is-hlapiti, vapiti
~epi	~epljah	~epi	~epivši	za-čepiti, cijepiti, lijepiti, za-slijepiti, o-krijepiti
~ipi	~ipljah	~ipi	~ipivši	đipiti, škripiti
opi		opij	opivši	
škopi	škopljah	škopi	škopivši	
pre-klopi		pre-klopi	pre-klopivši	
škropi	škropljah	škropi	škropivši	
topi	topljah	topi	topivši	
~rpi	~rpljah	~rpi	~rpivši	is-crpiti, drpiti, krpiti
~upi	~upljah	~upi	~upivši	kupiti, lupiti, o-tupiti, stupiti
iz-ri		iz-rij	iz-rivši	
~ri	~rah	~ri	~rivši	~riti
kri	krijah	krij	krivši	
~si	~šah	~si	~sivši	viele
ši	šijah	šij	šivši	
iš-čaši		iš-čaši	iš-čašivši	

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
	2it	ipf	plašiti	plaše	plašila
	2it	pf	nad-mašiti	nad-maše	nad-mašila
	6itut	pf	rašiti	rašiju	rašila
	2it	pf	o-prašiti	o-praše	o-prašila
	2it	ipf	strašiti	straše	strašila
homC	2it		~ešiti	~eše	~ešila
	2it	pf	lišiti	liše	lišila
homC	2it		~ošiti	~ošē	~ošila
homC	2it		~ršiti	~rše	~ršila
homC	2it		~ušiti	~uše	~ušila
homC	2tit		~atiti	~ate	~atila
homC	2tit		~etiti	~ete	~etila
homC	2tit		~ititi	~ite	~itila
	2tit	ipf	pamtiti	pamte	pamtila
homC	2tit		~otiti	~ote	~otila
	2tit	pf	zaptiti	zapte	zaptila
	2tit	pf	na-prtiti	na-prte	na-prtila
	2tit-šć	pf/ipf	častiti	časte	častila
	2tit-št	pf	o-vlastiti	o-vlaste	o-vlastila
	2tit-št	ipf	mastiti	maste	mastila
	2tit-šć	pf	za-mastiti	za-maste	za-mastila
homC	2tit-št		~pastiti	~paste	~pastila

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
plaši	plašah	plaši	plašivši	
nad-maši		nad-maši	nad-mašivši	
raši		rašij	rašivši	
o-praši		o-praši	o-prašivši	
straši	strašah	straši	strašivši	
~eši	~ešah	~eši	~ešivši	smiješiti, riješiti, griješiti, kostriješiti, po-spešiti, tješiti
liši		liši	lišivši	
~oši	~ošah	~oši	~ošivši	trošiti, pustošiti
~rši	~ršah	~rši	~ršivši	kršiti, ras-pršiti, stršiti, vršiti
~uši	~ušah	~uši	~ušivši	bušiti, gušiti, njušiti, pušiti, rušiti, sušiti
~ati	~aćah	~ati	~ativši	viele
~eti	~ećah	~eti	~etivši	primijetiti, prijetiti, pro-svijetiti, sjetiti, posjetiti, o-sujetiti, o-pametiti, po-remetiti, teretiti, o-štetiti, o-svetiti
~iti	~ićah	~iti	~itivši	hititi, kititi, raz-golititi, pod-mititi, za-sititi, štititi
pamti	pamćah	pamti	pamtivši	
~oti	~očah	~oti	~otivši	kotiti, sramotiti, krotiti
zapti		zapti	zaptivši	
na-prti		na-prti	na-prtivši	
časti	čašćah	časti	častivši	
o-vlasti		o-vlasti	o-vlastivši	
masti	maštah	masti	mastivši	
za-masti		za-masti	za-mastivši	
~pasti	~paštah	~pasti	~pastivši	za-prepastiti, u-propastiti

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
homC	2tit-št		~estiti	~este	~estila
	2tit-šć	ipf	čistiti	čiste	čistila
	2tit-št	ipf	koristiti	koriste	koristila
	2tit-šć	ipf	gostiti	goste	gostila
	2tit-šć	ipf	žalostiti	žaloste	žalostila
	2tit-št	pf	pre-mostiti	pre-moste	pre-mostila
	2tit-št	pf	oprostiti	oproste	oprostila
	2tit-št	pf/ipf	krstiti	krste	krstila
	2tit-št	pf	u-vrstiti	u-vrste	u-vrstila
	2tit-šć	pf	pri-čvrstiti	pri-čvrste	pri-čvrstila
	2tit-št	pf	pustiti	puste	pustila
homC	2it		~štiti	~šte	~štila
homC	2it		~utiti	~ute	~utila
basC	2bitmitpitvit		~viti	~ve	~vila
	6itut	ipf	viti	viju	vila
	6itut	pf	obaviti	obaviju	obavila
homC	2gitzit		~ziti	~ze	~zila
homC	2it		~žiti	~že	~žila
	7st-s	pf	pasti	pasu	pasla
	7st-dn	ipf	pasti	padnu	pala
	7st-st	ipf	rasti	rastu	rasla
	7st-d	ipf	krasti	kradu	krala
	7st-t	ipf	cvasti	cvatu	cvala
	7st-d	ipf	jesti	jedu	jela
	7st-d	pf	djesti	djedu	djela
	7st-dn	pf	sjesti	sjednu	sjela
	7st-t	ipf	plesti	pletu	plela
	7st-t	ipf	mesti	metu	mela
	7st-d	ipf	presti	predu	prela
	7st-tn	pf	sresti	sretnu	srela
	7st-s	ipf	tresti	tresu	tresla
	7st-z	ipf	vesti	vezu	vezla
	7st-d	pf	na-vesti	na-vedu	na-vela
	7st-z	ipf	gristi	grizu	grizla
	7st-d	pf	bosti	bodnu	bola
	7st-b	ipf	gresti	grebu	grebla
	7st-p	ipf	tepsti	tepu	tepla

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
~esti		~esti	~estivši	oba-vijestiti, o-nesvijestiti, na-mjestiti
	~eštah			
čisti	čiščah	čisti	čistivši	
koristi	korištah	koristi	koristivši	
gosti	goščah	gosti	gostivši	
žalosti	žaloščah	žalosti	žalostivši	
pre-mosti		pre-mosti	pre-mostivši	
oprosti		oprosti	oprostivši	
krsti	krštah	krsti	krstivši	
u-vrsti		u-vrsti	u-vrstivši	
pri-čvrsti		pri-čvrsti	pri-čvrstivši	
pusti		pusti	pustivši	
~šti		~šti	~štivši	laštiti, u-skladištiti, po-ništititi, pištiti, na-mrštiti, priuštiti, ljuštiti
	~štah			
~uti		~uti	~utivši	slutiti, ljutiti, o-šamutiti, s-putiti
	~utah			
~vi	~vljah	~vi	~vivši	~viti
vi	vijah	vij	vivši	
obavi		obavij	obavivši	
~zi	~žah	~zi	~zivši	viele
~ži	~žah	~ži	~živši	viele
pase		pasi	pasavši	
pade	padah	padni	pavši	
raste	rastah	rasti	rastavši	
krade	kradah	kradi	kradavši	
cvate	cvatah	cvati	cvatavši	
jede	jedah	jedi	jedavši	
djede		djedi	djedavši	
sjede		sjedni	sjevši	
plete	pletah	pleti	pletavši	
mete	metah	meti	metavši	
prede	predah	predi	predavši	
srete		sretni	srevši	
trese	tresah	tresi	tresavši	
veze	vezah	vezi	vezavši	
na-vede		na-vedi	na-vedavši	
grize	grizah	grizi	grizavši	
bode		bodi	bodavši	
grebe	grebah	grebi	grebavši	
tepe	tepah	tepi	tepavši	

Cluster	Flexionstyp	Asp.	Infinitiv	Präsens 3. Pl.	Prät. Sg. fem.
	7st-b	ipf	zepsti	zebu	zebla
	7st-p	pf	is-crpsti	is-crpu	is-crpla
	7st-b	ipf	dupsti	dubu	dubla
	7st-z	ipf	musti	muzu	muzla
	6itut	pf	obuti	obuju	obula
	6itut	pf/ipf	čuti	čuju	čula
homC	3ut		~nuti	~nu	~nula
	3ut-p	pf	oba-suti	oba-spu	oba-sula
	6itut	pf	izuti	izuju	izula

Aorist 3. Sg.	Imperfekt 1. Sg.	Imperativ Sg.	Adverbialpart. Prät.	Kommentare
zebe	zebah	zebi	zebavši	
is-crpe		is-crpi	is-crpavši	
dube	dubah	dubi	dubavši	
muze	muzah	muzi	muzavši	
obu		obuj	obuvši	
ču	čujah	čuj	čuvši	
~nu		~ni	~nuvši	viele
oba-su		oba-spi	oba-suvši	
izu		izuj	izuvši	

Teil IV.

Bibliographie

Inhalt Teil IV

1. ALLGEMEINE BIBLIOGRAPHIE.....	IV.3
2. KROATISCHE BIBLIOGRAPHIE.....	IV.5

1. Allgemeine Bibliographie

Alpar, Paul; Niedereichholz, Joachim: *Data Mining im praktischen Einsatz*. Braunschweig / Wiesbaden 2000.

Bußmann, Hadumod (ed.): *Lexikon der Sprachwissenschaft*. Stuttgart³2002.

Ester, Martin; Sander, Jörg: *Knowledge discovery in databases. Techniken und Anwendungen*. Berlin 2000.

Fayyad, Usama (ed.): *Advances in Knowledge Discovery and Data Mining*. Menlo Park CA 1996.

Holl, Alfred; Maroldo, Sara; Urban Reinhard: *The inflectional morphologies of the Swedish noun, the Swedish verb and the English verb. Reverse dictionaries based upon data mining methods*. Växjö 2007 [= Mathematical modelling in physics, engineering and cognitive sciences, vol. 12].

Holl, Alfred; Pavlidis, Stilianos; Urban, Reinhard: *Rückläufiges Wörterbuch zur alt- und neugriechischen Verbalmorphologie. Aufbereitung mit Datenanalyse-Verfahren der Informatik (Data Mining)*. Regensburg 2006 [= *Studia et exempla linguistica et philologica*, Series V: Lexica, Tom. 5].

Holl, Alfred; Behrschmidt, André; Kühn, Alexander: *Rückläufige Register zur russischen und deutschen Verbalmorphologie. Aufbereitung mit Datenanalyse-Verfahren der Informatik (Data Mining)*. Regensburg 2004 [= *Studia et exempla linguistica et philologica*, Series V: Lexica, Tom. 4].

Holl, Alfred: "Datenanalyseverfahren der Informatik (Data Mining) als Grundlage einer didaktischen Darstellung der französischen Verbalmorphologie". In: Bernhard, Gerald; Kattenbusch, Dieter; Stein, Peter (ed.): *Namen und Wörter. Festschrift Josef Felixberger zum 65. Geburtstag*. Regensburg 2003, 107-119.

Holl, Alfred: "Licht und Schatten von Analogieschlüssen auf der Basis rückläufiger Ähnlichkeit in der Verbalmorphologie romanischer und germanischer Sprachen". In: Heinemann, Sabine; Bernhard, Gerald; Kattenbusch, Dieter (ed.): *Roma et Romania. Festschrift Gerhard Ernst zum 65. Geburtstag*. Tübingen 2002, 152-167.

Holl, Alfred: "The inflectional morphology of the Swedish verb with respect to reverse order: analogy, pattern verbs and their key forms". *Arkiv för nordisk filologi* 116 (2001), 193-220.

Holl, Alfred: *Romanische Verbmorphologie und relationentheoretische mathematische Linguistik. Axiomatisierung und algorithmische Anwendung des klassischen Wort und Paradigma-Modells*. Tübingen 1988 [= Linguistische Arbeiten 216].

Householder, Fred W.: "Descriptive analysis of Latin declension". *Word* 3 (1947) 48-58.

Kempgen, Sebastian: *Grammatik der russischen Verben*. Wiesbaden 1989.

Kluge, Friedrich: *Etymologisches Wörterbuch der deutschen Sprache*. Berlin²⁴2002.

Krahl, Daniela; Windheuser, Ulrich; Zick, Friedrich-Karl: *Data Mining. Einsatz in der Praxis*. Bonn 1998.

Kruse, Rudolf; Borgelt, Christian: "Suche im Datendschungel". *Spektrum der Wissenschaft* 2002: 11, 80-81.

Langendorf, Dieter: *L'art de conjuguer. Le Nouveau Bescherelle: Dictionnaire de douze mille verbes*. Frankfurt 1986.

Lühr, Rosemarie: *Neuhochdeutsch. Eine Einführung in die Sprachwissenschaft*. München⁶2000.

Mateo, Francis: *Collection Bescherelle. El arte de conjugar en espanol. Diccionario de 12 000 verbos*. Paris 1995.

Mater, Erich: *Rückläufiges Wörterbuch der deutschen Gegenwartssprache*. Leipzig 1965.

Müller, Sonja: *Zur Systematik morphologischer Paradigmen. Die deutschen Verben in der Computermorphologie*. Heidelberg 1992.

Muthmann, Gustav: *Rückläufiges deutsches Wörterbuch*. Tübingen 2001.

Schweiger, Fritz: "Review of Holl 1988". *Yearbook of Morphology* 3(1990) 238-240.

Simmler, Franz: *Morphologie des Deutschen: Flexions- und Wortbildungsmorphologie*. Berlin 1998.

Trost, Klaus: "Agentivitätsstruktur, Postfigierung und Reflexivität im Russischen". *Anzeiger für slavische Philologie* 24 (1996) 57-76.

Trost, Klaus: "Das System der Durativitätsarten, der Aktionsarten und des Aspekts im Russischen". *Anzeiger für slavische Philologie* 30 (2002) 21-55, 31 (2003) 199-232.

Ulrich, Winfried: *Wörterbuch linguistische Grundbegriffe*. Berlin ⁵2002

Weermann, Eva Maria: *Pons Verbtabelle Deutsch*. Stuttgart 2001.

Zaliznjak, Andrej A.: *Grammatičeskij slovar' russkogo jazyka: slovoizmenenie*. Moskau ⁴2003 = ³1987 = ¹1977 [= Зализняк, Андрей А.: *Грамматический словарь русского языка: словоизменение*].

2. Kroatische Bibliographie

Anić, Vladimir: *Veliki rječnik hrvatskoga jezika*. Zagreb 2006.

Babić, Stjepan; Božidar, Finka; Moguš, Milan: *Hrvatski pravopis*. Zagreb 1994.

Engel, Ulrich; Mrazović, Pavica (ed.): *Kontrastive Grammatik Deutsch-Serbokroatisch*. Bd. 1. München 1986.

Gubišić, Vinko: *Croatian grammar*. Zagreb 1995.

Kunzmann-Müller, Barbara: *Grammatikhandbuch des Kroatischen unter Einschluss des Serbischen*. Frankfurt ³2002 [= Heidelberger Publikationen zur Slavistik: A, Linguistische Reihe; Tom. 7].

Leitner, Hans: *Deutsch-kroatisches Wörterbuch: Verben im Kontext*. Zagreb 1998.

Panzer, Baldur; Ressel, Svetlana: *Valenzwörterbuch kroatischer und serbischer Verben*. Frankfurt 2001 [= Heidelberger Publikationen zur Slavistik: A, Linguistische Reihe; Tom. 11].

Težak, Stjepko; Babić, Stjepan: *Gramatika hrvatskoga jezika. Priručnik za osnovno jezično obrazovanje*. Zagreb ¹⁵2005.

Anhang. Beschreibung des Inhalts der CD

Die Datei **HR_Verbregister.xls** enthält das Register der kroatischen Verben, das sich in III.4 findet.

Die Datei **HR_Flexionstypen.xls** enthalten die Flexionstypen aller kroatischen Verben. Der Inhalt dieser Datei ist aus Abbildung III.7 ersichtlich.

Systemvoraussetzungen

Pentium PC mit mind. 1 GHz Prozessor, 128 MB RAM, Win2000, XP, CD-ROM-Laufwerk, MS Office XP.