Tel Aviv University
The Lester & Sally Entin Faculty of Humanities
The Shirley & Leslie Porter School of Cultural Studies

Mapping Hebrew Dative Constructions

Thesis submitted for the degree "Doctor of Philosophy"

By

Elitzur Dattner

Submitted to the Senate of Tel Aviv University

August, 2014

Tel Aviv University
The Lester & Sally Entin Faculty of Humanities
The Shirley & Leslie Porter School of Cultural Studies

Mapping Hebrew Dative Constructions

Thesis submitted for the degree "Doctor of Philosophy"

By

Elitzur Dattner

Submitted to the Senate of Tel Aviv University

August, 2014

This work was carried out under the supervision of
Professor Mira Ariel

Contents

Acknowledgments **i**

Abstract **iii**

1 Introduction **1**
1.1 Too many definitions 2
1.2 Too many syntactic environments 3
1.3 The solution 4

2 Theoretical background, Data and Method **9**
2.1 Theoretical framework 9
2.1.1 A Construction Grammar approach to Argument Structure Constructions 9
2.1.2 A semantic approach to Grammar: Types, roles, and relations 11
2.1.2.1 Semantic types 11
2.1.2.2 Semantic roles and Core arguments 13
2.1.3 Transitivity and Affectedness 14
2.1.4 Usage-based theories of language and cognition 16
2.2 Data 18
2.3 Method 20

3 Dative participant roles in Hebrew **33**
3.1 Addressee 36
3.2 Experiencer 37
3.3 Evaluative Reference Point 39
3.4 Human Endpoint 41
3.5 Recipient 46
3.6 Discourse Marker 49
3.7 Affectee and Possessive Dative 50

3.8 Ethical Dative . . . 53
3.9 Summary . . . 54

4 Corpus analysis **55**
4.1 From Dative functions to Argument Structure Constructions . . . 55
4.2 Converging functions: Discourse profile constructions . . . 61
4.2.1 Correspondence Analysis . . . 62
4.2.2 Multiple Correspondence Analysis . . . 65
4.2.2.1 Type of predicate . . . 70
4.2.2.2 Number of arguments in the clause . . . 81
4.2.2.3 Type of the non-Dative argument . . . 83
4.2.2.4 Word order . . . 85
4.2.2.5 Lexical category of the predicate . . . 87
4.2.2.6 Verbal paradigm (Binyan) . . . 89
4.2.2.7 Mode . . . 94
4.2.2.8 Person of the Dative-marked participant . . . 94
4.2.2.9 Agentivity of the A participant . . . 94
4.2.2.10 Individuation of the O participant . . . 94
4.2.3 Converging Dative functions . . . 100

5 Cluster analysis **105**
5.1 Introduction to clustering . . . 106
5.2 Cluster One: Extended transitive Discourse Profile Construction . 111
5.2.1 Affective situations . . . 113
5.2.2 Extension of an affective event . . . 115
5.3 Cluster Two . . . 116
5.4 Cluster Three . . . 119
5.5 Cluster Four . . . 122
5.6 Cluster Five . . . 124

6 Conclusions **127**

Bibliography **131**

A Predicate list **139**

B MCA summary **169**

C HCPC description **177**
C.1 Description of the clusters according to individuals: central and unique exemplars 177
C.2 Description of clusters according to variables 178

Acknowledgments

Writing a doctoral dissertation may seem to be a one-man-show, but it is far from it. On every step of this very long journey I have leaned on a dear group of people for mental support, reaching for their insights in every interaction.

First and foremost, I would like to thank my supervisor, Prof. Mira Ariel, who was my mentor for the last ten years, since my third year of undergraduate studies. Her 2005 seminar, Pragmatics and Grammar, opened my eyes to new approaches to language, different from the ones I was used to. Mira helped me see exciting places where I could find answers for the questions I was asking at the time. Ever since then, Mira has supported me, mentally and with a fellowship. She walked with me as I discovered yet newer places to look for answers, patient with my slow writing pace, and accepting out-of-the-box thinking. To this day, I consider her to be a totem pole of alternative to the mainstream, allowing me to break through and find my own way in the sea of linguistic thought. For that I am most grateful.

Prof. Dorit Ravid, not an official part of my PhD in any way, was nonetheless very important to it. The atmosphere this dissertation was written in has a lot to do with my being around her. Prof. Ravid hired me as a team member in an inter-disciplinary project, and I was amazed by the width and depth of her knowledge, by her quick, associative thinking, and by the scope of the resources that inspire her. The project ended, but I remained around Prof. Ravid in many constellations, participating in her wonderful research group, working with her and Dr. Amalia Bar-On on their co-edited book, and helping some of her students with their own research. While Prof. Ravid's research and mine do not officially overlap, I was inspired by her, both theoretically and philosophically.

My parents, Naomi and Reuven Dattner, deserve every kind and warm word I can think of. They supported my growing family through all this time, and practically enabled this endeavor. My father and my mother are both intellectuals, of different kinds. My father is an autodidact teen-age spirited criticizer of the obvious with a sharp perception of the world and its wrongs. My mother is a mediator, in practice and in thought, that acknowledges the complexity of relations and interactions. I was raised between these two poles. They raised me to think, to

criticize, to argue, and to accept. Everything I have learned I owe them, because they taught me to learn.

Dr. Itai Dattner, my brother, an insightful sounding board for new ideas, has helped me numerous times to clarify my thoughts to myself. Whether in a 45-minute Skype call to the Netherlands or a three-hour meeting at his Haifa office, a five minute good-morning call on the way to work or through email correspondence, his meticulous, methodological way of thinking forced me to crystallize every argument I laid before him.

My parents-in-law, Zehava and Ilan Laniado, have helped us in these times of need, and allowed me to keep doing whatever it is I was doing with a less troubled mind.

Finally, Lee Dattner, my wife, a lighthouse in stormy nights, a rock in rapid waters, my love who knows how to take me by the hand to the right drawer in my mind, where I could find peace. I owe you so much. For your patience, for your understanding, for your endurance, for your optimism, for your love.

Abstract

As a grammatical case, the essential function of the Dative (crosslinguistically) is to describe a particular relation between a state of affairs (an event or a state) to a referent. That is, it is a tool provided by the language for the speaker to portray how a specific situation is related to an entity in the world. If so, what is unique about the Dative? What is the difference between the Dative and other cases, other tools the language provides for describing relations between situations and referents, like the Accusative, or the Nominative, for instance, that also mark relations between X and Y? The Dative, as its syntactic status in many languages (an Indirect Object) suggests, marks an *indirect relation*. Such an indirect relation might be manifested as an indirect or partially (as opposed to completely) affected participant in an event, or as a secondary, non-inherent participant in a situation. This type of indirect relation is usually related to cognitive, rather than physical, involvement. This quality of the Dative, I argue in the present research, is unique to the marking of *human patients.* Human patients tend to be marked as non-canonical direct objects (Haspelmath, 2001). That is, their patient status is marked relative to an inanimate theme: human patients are marked as going through an incomplete change of state.

There are different kinds of indirect relations, all marked by the same morphological Dative form in Hebrew. A Recipient of an object, for instance, is an indirect participant of a transitive motion event in which an Agent moves a Theme from point a to point b. An Evaluator of a situation is another indirect participant in the situation itself. This uniform marking of different types of participants raises two questions. First, how are these different types of relations related to each other (see Boneh and Bar-Asher Siegal (2014) for a recent attempt to answer this question). Second, how are these types differentiated by the speaker/hearer, if differentiated at all? That is, when a speaker utters a clause with a Dative marked Indirect Object, what interpretation strategy is expected from the hearer in order to interpret the right type of Dative? It is the goal of this study to answer these questions. I argue that there are four basic types of Dative functions in Hebrew, each related to a particular discourse pattern, and they are all related to each other

on the basis of an innovative Affectedness Scale. Thus, the discourse pattern (and specifically, the level of discursive transitivity and subjectivity of the clause) is what points the hearer to the right interpretation of the Dative at hand.

This assumption does not necessarily concur with previous accounts of the Dative, in Hebrew and cross-linguistically. Traditionally, Dative functions are described and explained regardless of their context, whether syntactic or discursive. That is, the explanation of each Dative function is limited to the nature of the relation between the Dative referent and the Direct Object, for example (asking whether there are possession relations between the two or not), or to the particular semantic role played by the Dative-marked participant in the event (whether it is a Deprivee or a Malefactive, for example). The present study aims at describing and explaining the Dative in a wider context. Instead of having the Dative morpheme itself as the object of research, I aim to account for Dative constructions.

This account of the Hebrew Dative is derived from a quantitative, Usage-Based approach to language. Through a corpus research, I have considered not only syntactic or semantic features in isolation but taken into account multiple parameters from multiple sources of linguistic and extra-linguistic information. The corpus serving as a database for the present research is an approximately 1,760,000 words corpus of spoken Hebrew. The corpus is a collection of transcriptions of 198 meetings of committees of the Knesset, the Israeli parliament, composed of multiple registers of language, both formal and colloquial. From this corpus, all occurrences of a Dative (*le-*) marked pronoun were extracted, resulting with 16,575 tokens of Dative uses. Each token was then coded by hand for 17 parameters.

The analysis presented in this thesis is based on multivariate exploratory statistics, and specifically on Multiple Correspondence Analysis and Hierarchical Clustering on Principal Components. Multiple Correspondence Analysis is an exploratory statistical techniques that reveals frequency-based associations in the data, visualizing these associations in the form of a map. Highly associated forms, for instance, appear closer on the map than forms with no association. The structure of the data and the grouping of tokens within it was described using Hierarchical Clustering on Principal Components. Hierarchical Clustering on Principal Components (HCPC) is a statistical tool that partitions the data into sets of similar individuals. It is a complementary tool of Multiple Correspondence Analysis, used for graphically representing similarities or correlations, and for providing statistically defined clustering of tokens in the data.

The present study proposes a new theoretical concept, and defines a usage profile as a Discourse Profile Construction: a conventional pairing of multiple sources of information with a particular construal of the world. Assuming an

exemplar-based cognitive representation of language, and based on the findings from the multivariate exploratory statistics, four Discourse Profile Constructions emerge from the corpus:

1. The Extended Transitive Discourse Profile Construction,
2. The Human Endpoint Discourse Profile Construction,
3. The Extended Intransitive Discourse Profile Construction,
4. The Evaluative Reference Point Discourse Profile Construction.

These four Discourse Profile Constructions represent four sets of Dative tokens in the corpus, each of which is characterized by a different usage pattern. The present research shows that from a usage point of view there is little importance to the subjective, predicate-specific interpretation of the participant role itself. That is, the substantial differences are shown to exist between Discourse Profile Constructions representing sets of Dative functions converged under a single construal of the world, rather than between locally defined participant roles such as an Experiencer or a Recipient. These Discourse Profile Constructions are the basis for comparison between different tokens of use for both the interpretation of the Dative and the expansion of the cognitive category of each construction.

1 Introduction

As a grammatical case, the essential function of the Dative (crosslinguistically) is to describe a particular relation between a state of affairs (an event or a state) to a referent. That is, it is a tool provided by the language for the speaker to portray how a specific situation is related to an entity in the world. If so, what is unique about the Dative? What is the difference between the Dative and other cases, other tools the language provides for describing relations between situations and referents, like the Accusative, or the Nominative, for instance, that mark relations as well? The Dative, as its syntactic status in many languages (an Indirect Object) suggests, marks an *indirect relation*. Such an indirect relation might be manifested as an indirect or partial (as opposed to complete) affected participant in an event, or as a secondary, non-inherent participant in a situation. This type of indirect relation is usually related to cognitive, rather than physical, involvement. This quality of the Dative, I argue in the present research, is unique to the marking of *human patients.* Human patients tend to be marked as non-canonical direct objects (Haspelmath, 2001). That is, their patient status is marked relative to an inanimate theme: human patients are marked as going through an incomplete change of state.

There are different kinds of indirect relations, all marked by the same morphological Dative form. A Recipient of an object, for instance, is an indirect participant of a transitive motion event in which an Agent moves a Theme from point a to point b. An Evaluator of a situation is another indirect participant in

the situation itself. This uniform marking of different types of participants raises two questions. First, how are these different types of relations related to each other (see Boneh and Bar-Asher Siegal (2014) for a recent attempt to answer this question). Second, how are these types differentiated by the speaker/hearer, if differentiated at all? That is, when a speaker utters a clause with a Dative marked Indirect Object, what interpretation strategy is expected from the hearer in order to interpret the right type of Dative? It is the goal of this study to answer these questions. I argue that there are four basic types of Dative functions in Hebrew, each related to a particular discourse pattern, and they are all related to each other on the basis of a particular Affectedness Scale. Thus, the discourse pattern (and specifically, the level of discursive transitivity and subjectivity of the clause) is what points the hearer to the right interpretation of the Dative at hand.

This assumption does not necessarily concur with previous accounts of the Dative, in Hebrew and crosslinguistically. Traditionally, Dative functions are described and explained regardless of their context, be it syntactic or discursive. That is, the explanation of each Dative function is limited to the nature of the relation between the Dative referent and the Direct Object, for instance (asking whether there are possession relations between the two or not), or to the particular semantic role played by the Dative-marked participant in the event (whether it is a Deprivee or a Malefactive, for example). The present study aims at describing and explaining the Dative in a wider context. Instead of having the Dative morpheme itself as the object of research, I aim to account for Dative constructions.

In the following section I present some of the well known accounts of the Dative. Next, I discuss another source of problems in providing an account of the Dative, namely, the multiple, unrelated syntactic environments it appears in. Finally I propose my solution, to be developed in the rest of the study.

1.1 Too many definitions

The dative case has several recurrent functions in language after language, together with a set of language-specific functions. The typical dative functions are direction, recipient, experiencer, purpose, possessor, and beneficiary (Haspelmath, 2003).

Dative marked participants relate to certain types of verbs crosslinguistically: Possession, existence, psychological states, physiological states, visual or auditory perceptions, modal states of necessity and wanting, modal states of potentiality, and uncontrolled events (Shibatani, 2001).

The Dative case has been given many definitions over the years, from different

points of view within different theoretical schools. Here are only few of these points of view. Berman (1982), discussing the Hebrew Dative, lists several types of uses of the Dative, showing that basically they all share a single quality, and concluding that the Dative in Hebrew marks an Affectee. Similarly, Dabrowska (1997), working within a cognitive linguistic framework, defines the Polish Dative as a Target Person, which is: "an individual who is perceived as affected by a change, activity, or state in his or her personal sphere" (Dabrowska, 1994, p. 110).

Van Belle and Langendonck (1996); Van Langendonck and Van Belle (1998) are two major sources for Dative studies in a typological perspective, providing many different definitions and describing various behavioural properties of the Dative in many languages. Considering the amount of research presented in these two seminal volumes, no single conclusion can be drawn with regard to a unified definition of 'The Dative.' Also within a typological perspective, Haspelmath (2003) defines the Dative using a semantic map which revolves around the Recipient as the prototypical Dative function. The Affectee, or Beneficiary, is an extension of the prototype, according to Haspelmath. In Japanese, Kishimoto (2010) concludes that the semantic basis of Dative marking is possession.

Considering some uses of the Hebrew Dative, Borer and Grodzinsky (1986) discuss the argumenthood status of the Dative. Contrary to other accounts of the Hebrew Dative, such as Berman (1982), for example, no unified description of the Dative is suggested. Rather, the Dative is used as a tool for providing an argument for a particular syntactic structure.

Other researchers have been focused on particular uses of the Dative, aiming at typological conclusions, syntactic-semantic ones, or lexical-semantics generalizations. For example, see Sridhar (1979); Blume (1998); Šaríc (2002); Cuervo (2003); Amritavalli (2004); Francez (2006); Hole et al. (2006); Levin (2008); Ariel et al. (In progress).

One crucial problem with such a vast amount of accounts, taking into consideration that very few of these works are based on corpus data, is the fluid nature of grammatical judgements most of these works are based on. As a consequence, different judgements and different theoretical frameworks lead to utterly different descriptions of basically the same phenomena. Yet this is only the first problem one has to tackle when attempting an account of the Dative. The second problem is described in the next section.

1.2 Too many syntactic environments

Consider the following examples:

(1) a. *shalaxti lo mixtav she-oto ash'ir laxem kan.*
I.sent to.him letter that-it I.leave to.you here.
'I sent him a letter which I leave here for you.'
(10,141)

b. *ani mash'ir lahem lehacig et ha-ta'arix shelahem.*
I leave to.them to present ACC the-date their.
'I leave it for them to say when they'll do it.'
(9,979)

c. *bou nizkor she-nish'aru laxem reservot me-ha-shana sh-avra.*
come remember that-were.left to.you extras from-the-year that-pass.
'Let's not forget that you have some extras left from last year.'
(10,681)

The problem raised by these examples is a disambiguation problem. While they all share one and the same predicate they have different argument structures, and the Dative interpretation is different as well. In (1a) it is a Recipient or a Beneficiary, in (1b) it is some type of an enabled person, while in (1c) it might be termed a Possessor. One approach to solve this problem is assuming Argument Structure Constructions and metaphoric extensions of which (a solution adopted in Goldberg (1995), for example). However, in the course of the study we will see that this solution is not an adequate one, since one Argument Structure Construction can accommodate more than one function. Another way to solve this problem is to search for a wider context of cues for interpretation, rather than considering syntax or semantics in isolation. This approach is introduced in the next section.

1.3 The solution

The present work presents a Usage-Based approach to the interpretation of Hebrew Dative clauses. This approach is based on a multivariate statistics analysis of a natural language corpus, utilizing two complementary tools for uncovering patterns in large data, and propose the concept of Discourse Profile Constructions: an emergent form-function conventional correlation that consists of multiple sources of formal and functional information. Taking into consideration multiple aspects of usage, both linguistic and extra-linguistic, I develop an Affectedness Scale according to which a Dative clause is interpreted.

Working within an Exemplar-based representation framework, I define patterns of usage as *Discourse Profile Constructions*. Comparing the Discourse Pro-

file Construction account of the Hebrew Dative with the more traditional function/participant role account, I show that from a usage point of view there is very little significance to the subjective decision about particular participant roles. I show that each Discourse Profile Construction accommodates sets of participant roles with similar usage profiles. The particular usage profile constitutes the basis for comparison between different tokens of use for interpretation, and for the expansion of the cognitive category of each construction, through similarity between exemplars.

The meaning of the Hebrew Dative on its own is abstract, determined only through the Discourse Profile Construction it is used in. I propose that the Dative-marked participant is a reference point for an event or a state conveyed by the predicate in the utterance. This reference point subjectifies the situation's construal by profiling its affectee, its audience, its experiencer, or an anchor for evaluation. Thus, if a clause displays high transitivity, it is more likely to convey an effect on the Dative-marked participant mediated by a manipulation of a different object. If a clause displays an intermediate level of transitivity, it is likely to convey an effect on the Dative-marked participant directly inflicted upon the human by the Agent. It is a partial effect, hence the (1) dative marking, and (2) intermediate transitivity. If a clause displays low transitivity it is likely to profile the Dative-marked participant as an Experiencer or Reference point for evaluation of a situation.

In some cases it is difficult to tell apart one interpretation from another, for instance the differences between the Affectee and the Possessive Dative, or between the Experiencer and the Evaluative reference point. While these fine-grained semantic differences between closely related participant roles are worthy topics for research in many cases, the current study suggests that from a usage-based point of view these different roles are sub-varieties of a single function. That is, from a usage-based point of view, I claim that we should look at functions converged under a single construction.

Current accounts of the Hebrew Dative accurately describe the different Dative functions. However, most of them either approach the problem from a narrow semantic-syntactic perspective, or account for each function in isolation without considering its relation to the other functions served by the Hebrew Dative. Moreover, as most of these accounts are not corpus-based, they remain subjective, depending on different intuitive grammaticality judgements. Conversely, through a corpus analysis of the Dative in Hebrew, I ask what is the organizing principle of the Hebrew Dative that both categorizes different uses and functions into a single cognitive category, and differentiates each function from the others based on its usage profile. I suggest an organizing principle that takes the gradient status

of grammaticality into consideration, claiming that there are three major Dative functions, organized in Discourse Profile Constructions according to Transitivity and Subjectivity parameters. Each of these Discourse Profile Constructions unifies closely related types of the Dative into a single discourse pattern. The differences between functions within the same Discourse Profile Construction are accounted for in a quantitative way, showing that only by adopting a wide perspective of constructions rather than isolated functions, and of multiple parameters rather than focusing on the semantic-syntactic pairing, each Dative function can be objectively defined.

As noted above, the Dative interpretation is shown to be a factor of multiple transitivity parameters, including discursive ones. For instance, while the syntactic argument structure remains the same in (2)–(5), the different discourse profiles assigns different interpretation to the Dative-marked participant in each sentence. Note that examples with the same (or similar) discourse profile belong to the same usage pattern, whether they share a traditional interpretation or not. That is, it is my claim that the relevant information for interpreting a Dative clause is its usage pattern, or its Discourse Profile Construction, rather than its particular Dative-marked participant role. As an exemplification, let us consider the differences and similarities between the following sentences.

(2) ***lakxu lahem*** *et ha-dirot.*
took lahem ACC the-apartments.
'They took their apartments.'
(1,431)

(3) *tafsiku* ***levalbel lanu*** *et ha-moax.*
stop confusing **to.us** ACC the-brain.
'Don't talk nonsense.'
(1,902)

(4) *she-lo* ***ta'ase lanu*** *gader hafrada.*
that-not you.will.do **to.us** fence segregation.
'Just don't build a segregation fence (on us).'
(749)

(5) *im ha-hearxut ha-zot lo tece la-poal ka'asher xas ve-xalila*
if the-preparation the-this not carried out when god forbid
kodxim lanu *neft b-a-yam ha-tixon azay [. . .]*
drill **to.us** oil in the Mediterranean then [. . .]
'If this preparation won't be carried out now that when we're drilling oil in the Mediterranean then [. . .]'
(996)

A Dative-function oriented interpretation would analyse these examples as manifesting different Dative participant roles: the Dative in (2) may be interpreted as either a Deprivee or a Possessive Dative; in (3) the Dative-marked participant can be interpreted as a Possessive Dative or an Experiencer; the Dative in (4) can be interpreted as a Malefactive or an Ethical Dative, and so is the case for (5). However, (2)–(5) share a usage pattern, and share a construal of a situation in which an Agent is affecting the Dative-marked participant through manipulating a third entity. This construal is related to high transitivity parameters, involving an interpretation of the Dative-marked participant as highly affected by an event. In contrary, consider (6)–(7):

(6) *hem yod'im lehagid kama dalaf* ***lahem***.
they know to say how much leak **to.them**.
'They can indicate how much is leaking.'
(6,549)

(7) *im ha-hore lo yaxol leshalem ve-xozeret* ***lo*** *ora'at keva [...]*
if the-parent not can to pay and-returns **to.him** standing order [...]
ha-irgun yoce nizok.
the-organization exits damaged.
'If the parent can't pay and his standing order bounces [...] the organization is damaged.'
(383)

The Dative-marked participant in (6)–(7) can be interpreted as either a Malefactive/Affectee (6) or a Possessive Dative (7). Although these interpretations were relevant for (2)–(3) as well, notice that the effect inflicted upon the Dative-marked participant is different in the two sets of sentences. And indeed, (6)–(7) belong to a different usage pattern that involves a stative situation construal in which the Dative-marked participant is mildly affected (or not affected at all). This construal is characterized with low transitivity parameters; the Dative-marked participant is a reference point for an evaluation or a description of a situation. While the "objective" situation could have been verbalized without a Dative (for example, using a Genitive for expressing possession), the mere motivation for using a Dative construction in these cases is to profile the Dative-marked participant as affected.

The structure of the dissertation is as follows. First, I introduce the theoretical framework this study is embedded in, together with the database for the research and the method used to analyse it. I then present the various functions the Hebrew Dative can serve according to the present corpus, showing that most of them are in fact predicate-specific. Next, in the main part of the work, I present a corpus

analysis of the Hebrew Dative, arguing that a unified account can be achieved only within a usage-based perspective that takes into account multifactorial usage patterns rather than isolated morphemes or particular semantic roles. Finally, I propose a new theoretical concept, the Discourse Profile Construction: a unique multifactorial usage pattern that groups a cluster of similar tokens together, conventionally linking them with a particular function. Thus, I answer the questions posed above with a coherent description of both the difference between different roles, and the commonalities between similar functions.

2

Theoretical background, Data and Method

2.1 Theoretical framework

2.1.1 A Construction Grammar approach to Argument Structure Constructions

The main framework in which I analyze the data in this study is Construction Grammar, as described and developed in the works of Fillmore and Kay (for example, Fillmore et al. (1988); Kay (1997); Kay and Fillmore (1999)), Goldberg (e.g. Goldberg (1995, 2003, 2005b, 2006); Goldberg and Jackendoff (2004)), Croft (2001) and Bybee and Eddington (2006); Bybee (2006) among many others.

Constructions , in Construction Grammar theories, are the basic units of grammar (Diessel, 2004): conventionalized form-meaning correspondences. The existence – or the necessity – of these form-meaning pairings is a subject for debate between different branches of the construction grammar framework. Goldberg states that constructions are "posited whenever there is evidence that speakers cannot predict some aspect of their form, function, or use from other knowledge of language (i.e. from other constructions already posited to exist)" (Goldberg, 2005a, p. 17). On the other hand, in usage-oriented models of construction gram-

mar constructions are a function of usage. That is, a particular linguistic sequence is a construction (theoretically and presumably cognitively) based on usage parameters such as frequency and discourse pattern, whether it is predictable from other constructions or not (Ariel (2008); Bybee (2010)).

A special case of constructions is the Argument Structure Construction (Goldberg, 1995): the form-meaning pairings in the language that concern schematic clausal expression, rather than fully or partially fixed constructions such as idioms or prefabs (Fillmore et al., 1988). The English Caused Motion Construction is an example for such a basic, schematic, Argument Structure Construction:

(1) Pat sneezed the napkin off the table.

Goldberg's inovative approach to Argument Structure Constructions shows that the meaning associated with (1) cannot be analyzed as coming from the particular verb (an intransitive non-motion related verb in this case); it has to be searched for elsewhere. Goldberg puts foreward the proposition that it is the specific Argument Structure Construction in (2) that bears the 'causing to move' meaning, such that any verb (with several functional limitations) inserted into the V slot of the construction will 'gain' such a meaning:

(2) Subj V Obj Obl
'X causes Y to move Z' (Goldberg, 1995, p. 3).

In the following chapters I argue for a broad definition of Argument Structure Constructions, taking into consideration multiple sources of information: lexical, morphological, syntactic, semantic, and discursive information. In the remaining of the section I present Goldberg's (1995) take on Argument Structure Constructions.

One claim about Argument Structure Constructions is of critical importance for the present analysis, that simple clause constructions are associated with construals of basic human experience (Goldberg, 1995, p. 5). In the account of Hebrew Dative Constructions advocated in the present study I show that on top of basic event types, we can approach Argument Structure Constructions from a usage-based, discursive point of view, thus broadening Argument Structure Constructions to include *discursive* functions as well, emphasizing the existence of basic discursive scenarios.

2.1.2 A semantic approach to Grammar: Types, roles, and relations

The present section lays out the fundamentals of the present approach to the coding process of the corpus. First I will present Dixon's (2005) concept of primary and secondary semantic types, exemplifying the semantic types found in the corpus. Then I go on to present Dixon's (2006) concept of core semantic roles, defining the Hebrew Dative as an *Extension to core*. These notions, of semantic types and semantic roles, will serve us later in accounting for the different functions the Hebrew Dative manifests according to the present corpus.

2.1.2.1 Semantic types

The semantic type of each predicate in the corpus was analysed using a Hebrew dictionary (Even-Shoshan, 2003), and categorized into types along the lines of Dixon (2005). Dixon approaches semantic types in a straightforward manner. Each type of predicate is stripped down to its essence by treating it informally, focusing on the predicate's everyday meaning.

A semantic type, in Dixon's terminology, is a class of words that have a common meaning component. This is, of course, a categorization made on a semantic level. On the syntactic level, on the other hand, these semantic types can be grouped together as word classes (Dixon's concept of parts of speech) with common morphological and syntactic properties. Dixon argues that "[t]here is a relationship between semantic types and grammatical word classes" (ibid p. 9), such that each word class groups together some semantic types, in similar ways across languages. Thus, there might be a word class VERB in any given language which includes words belonging to semantic types such as MOTION, AFFECT, or SPEAKING.

Dixon categorizes predicates into semantic types, and into two sorts of verbal concepts: primary and secondary (Dixon, 2005, p. 96). Primary verbs directly refer to an activity or a state (and can stand on their own in a clause). Secondary verbs semantically modify some other verb. For instance, *hit* or *swim* are primary verbs, whereas *may, try*, or *let* are secondary. Dixon further divides these two sets into six subsets:

- Primary verbs:
  - Primary A: verbs with concrete participants as their semantic roles. For example, *natan*, 'give,' *salax*, 'forgive,' or *bitel*, 'cancel.' (see Appendix A for a complete list of predicates and semantic types).

  - Primary B: verbs with concrete participants that can be complemented by a clause either instead or in addition to the nominal complement, unlike Primary A verbs. For example, *amar*, 'say,' *katav*, 'write,' or *hidgish*, 'emphasize.'

- Secondary verbs:

  - Secondary A: verbs that have no independent semantic roles. Rather, they modify the meaning of another verb (the main verb of a complement clause), sharing its roles and syntactic relations. For example, *hispik*, 'be sufficient,' or *hifsik*, 'stop,' as in *hu hifsik lenagen be-gitara* 'He stopped playing the guitar,' in which the semantically main verb of the utterance is 'play,' rather than 'stopped.'
  - Secondary B: verbs that have one independent role (the subject of the Secondary-B verb) in addition to the roles of the verb in the complement clause. These verbs describe the subject's attitude towards some event or state. For example, *heemin*, 'believe,' *himtin*, 'wait,' or *cipa*, 'expect,' as in *ani mecape she-tagia maxar ba-zman* 'I expect you to arrive on time tomorrow.'
  - Secondary C: like Secondary-B verbs, Secondary C verbs introduce just one role (the subject of the Secondary verb) in addition to the roles of the verb in the complement clause. In Secondary C verbs, however, the subject actually plays a role in bringing about the event or state referred to in the complement clause's verb. For example, *azar*, 'help,' *garam*, 'cause,' or *hora*, 'instruct,' as in *hu hora lahem lehagia maxar ba-zman* 'He ordered them to arrive on time tomorrow.'
  - Secondary D: intransitive verbs that take a complement clause in subject slot. When these verbs take another role additional to those of the complement clause, it is some sort of stance taker, marked by the Dative in our case. For example, *nire*, 'seem,' or *nidme*, 'seem,' as in *nire li she-hu lo yagia maxar ba-zman* 'It seems to me that he won't be here on time tomorrow.'

Two types of adjectival predicates were found in the corpus as well. These are the Property type and the Value type. The Property type includes adjectives such as *barur*, 'clear,' *kashe*, 'hard,' and *nagish*, 'accessible.' The Value type includes adjectives such as *xashuv*, 'important,' *tov*, 'good,' and *nifla*, 'wonderful.'

Thus far we have reviewed the approach taken in accounting for the predicate of each clause in the corpus. The other aspect of Dixon's approach to grammar

which is relevant for the current research concerns the participants each clause is associated with – the semantic roles. This is the subject of the next section.

2.1.2.2 Semantic roles and Core arguments

Verbs refer either to an activity or to a state. In both activities and states there must be participants related to the situation. Dixon (2005) terms these participants Semantic Roles. Dixon's approach to semantic roles assumes that each semantic type of verbs is associated with a particular set of semantic roles, thus defining the semantic type. However, such an approach assumes that each semantic type is associated with a distinct set of roles. That is, for example, although both the transferred object in Transfer type verbs and the seen object in the See type are marked as Direct Object in Hebrew, semantically there is nothing in common between these roles. Dixon notes that in English "it is necessary to recognise forty or fifty semantic roles (ibid, p. 10)." There is no reason to assume otherwise for Hebrew.

From a syntactic point of view, however, there is a restricted number of Core Arguments. In this respect Dixon (2006) defines four clause types, based on the nature of their verbs and syntactic relations. The first clause type is the Intransitive clause, composed of an intransitive predicate and a single core argument, the Intransitive Subject (S). The second clause type is the Transitive clause, composed of a transitive predicate and two core arguments: a Transitive Subject (A) and a Transitive Object (O).

Many attempts have been made to solve the mapping problem from semantic roles to syntactic arguments (see, for instance, Dowty (1991); Grimshaw and Mester (1988); Reinhart (2003); Levin and Rappaport Hovav (2005), to name a few). In his writings, Dixon (1994, 2006) concludes that the argument whose referent initiates or controls the activity referred to in the predicate is mapped onto the A function, while the argument whose referent is affected by the activity is mapped onto the O function. Recall that according to Dixon's approach, which differentiate core arguments from semantic roles, these labels do not necessarily indicate the semantic role of their referents. That is, each core argument covers a range of semantic roles.

So far we have discussed three core arguments: an Intransitive Subject (S), a Transitive Subject (A), and a Transitive Object (O). Dixon further defines a fourth argument that can be added to both an Intransitive and a Transitive clauses (in Hebrew as well as in many other languages). This is the Extension to Core argument (E). In Hebrew, as in other languages, E is marked by the Dative. Thus, in such languages (including Hebrew), there are four types of clauses:

(3) Four clause types with core arguments (Dixon, 2006, p. 7)

(a)	intransitive	S		
(b)	extended intransitive	S		E
(c)	transitive	A	O	
(d)	extended transitive	A	O	E

An extended transitive clause in fact refers to clauses with extended transitive verbs, such as 'give' or 'show.' In Hebrew, such clauses will have the transferred object or the thing shown mapped onto the O function, and the Recipient or Addressee of the showing action mapped onto the E function (Dixon, 1994). Once again, we can see that a single core argument may cover a variety of semantic roles. The relations between the roles covered by the E core argument in Hebrew are the focus of the present research.

Summing up, the present section presented the grammatical aspect of the clauses we deal with in accounting for the Hebrew Dative, defining types of predicates and core arguments. The next section discusses another aspect of clauses in language, namely, Transitivity parameters.

2.1.3 Transitivity and Affectedness

Referring to the Hebrew Dative as an Extension to core transitive and intransitive clauses, we must consider a global aspect of transitivity. Hopper and Thompson (1980) present a notion of discourse-determined transitivity. Transitivity, as defined by Hopper and Thompson (1980), is gradient, composed of a configuration of ten parameters. These configurations define different degrees of transitivity, such that an action is transferred from one participant to another with different degrees of effectiveness or intensity (Hopper and Thompson, 1980, p. 252).

A high transitivity clause is a clause with two or more participants, conveying a telic, punctual action (rather than a state, or an atelic action) carried out by a volitional participant, high in Agency (the A participant). If the patient of the action, the O participant, is highly affected to the extent that it goes through a complete change of state, then the clause is higher in transitivity. If, moreover, the O referent is highly Individuated, such that it is human (or animate), concrete rather than abstract, singular, count, referential and definite, or marked by a proper noun rather than a common noun, the clause is yet higher in transitivity. Finally, affirmative clauses are higher in transitivity than negative ones, and so are clauses with a realis encoding of an event.

Transitivity has been found to play a crucial role in a particular type of Dative constructions in Hebrew (Dattner, 2008). Expanding on the scope of Dative con-

structions I account for in the present study, the corpus was coded for transitivity parameters, proposing transitivity as an organizing principle of the Hebrew Dative category as a whole, rather than of specific Dative constructions. See section 2.2 for a complete list of the transitivity parameters coded in the corpus.

One of the parameters differentiating transitive from intransitive clauses, as proposed in Hopper and Thompson (1980), for instance, is the affectedness of the O participant. It is commonly assumed that affectedness is not a dichotomous feature, but rather a continuum. Many have tried to define an affectedness scale, from Tsunoda (1985) to the very recent Beavers (2011). Most of these attempts have dealt with affectedness relative to a prototypical transitive event, and recent approaches to affectedness have concentrated on the minimal differences between different kinds of affected participants in a transitive event. This focus has lead the discussion towards defining and explaining different grades of a transitive clause, or a transitive predicate. That is, the focus of attention was not the participant itself, but rather what we can learn from its affectedness level about the lexical semantics of predicates or about transitivity in general.

Conversely, my interest in affectedness focuses on the affected *participant* and focuses on less than prototypical transitive clauses. And moreover, the nature of affectedness explored here is less direct to begin with. For instance, it would be hard to say that a *recipient* is directly and fully affected in the same way the patient of *break* is; the recipient is affected only in the sense that some object has been transferred into their control. Indeed, a fully affected entity is related to Core Transitive Verbs (Levin, 1999), while a less affected entity is related to Non Core verbs (and see also Haspelmath (2001) for an account of the non-canonical marking of non-prototypical objects).

Affectedness, then, is conceptualized here as a continuum spanning from a directly and physically affected entity, through a mentally affected entity, to an unaffected entity. Note that going down the affectedness scale we are drifting away from the prototypical transitive O (whether the direction is towards an indirect object or a transitive A is a question to be answered in later discussion).

The most affected entity is an entity that goes through some change of state as a direct result of a volitional Agent's action; this is the prototypical transitive O which undergoes a change of state caused by a transitive A. That is, a total effect is a resultative effect. This is the case, for example, with the O participant of *break* in which the impact of the action is a complete change of state. In the case of dative-marked participant roles, however, there is no clear (physical) change of state, yet they are still construed as affected entities in the conveyed state of affairs.

2.1.4 Usage-based theories of language and cognition

In recent years, usage-based approaches to language have developed in many directions, especially within the cognitive linguistics framework. A usage-based model of language assumes an "intimate relation between linguistic structures and instances of use of language." It is a model "in which the speaker's linguistic system is fundamentally grounded in 'usage events': instances of a speaker's producing and understanding language" (Kemmer and Barlow 2000:viii). Kemmer and Barlow (2000) note that a usage-based model emphasizes the role of frequency of instances and treats comprehension and production as integral to the linguistic system. In a usage-based theory linguistic representations are not fixed entities but rather emergent. And Elman (2009) finds out that "comprehenders' expectancies regarding the subcategorization frame in which a verb occurs is indeed sensitive to statistical patterns of usage that are associated not with the verb in general, but with the sense-specific usage of the verb".

Another important aspect of a usage-based model mentioned in Kemmer and Barlow (2000) is the intimate relation between usage, synchronic variation, and diachronic change (p. xviii). Frequency of use, together with certain discourse-based conditions, may cause a change in the language structure. As noted in Kemmer and Barlow (2000), "a dynamic, usage-based conception of the internal linguistic system provides a natural framework for understanding why variation and change exist in the first place, as well as for understanding the mechanisms that produce and propagate patterns of variation and change" (pp. xix–xx).

Ariel (2008), arguing for the role of discourse in determining use, presents the concept of the salient discourse pattern. It is the channel through which all external influences on language must go in order for a change to take place: "[i]t is ad hoc discourse-related considerations which prompt speakers to opt for one solution over another" (Ariel 2008:178). A particular cognitive distinction would become useful or not based on its discourse function. Ariel indicates that a novel meaning can evolve into a linguistic meaning only through "a motivated, transparent association between form and function, made available by heavy reliance on a supportive context" (Ariel 2008:182). These 'intimate relations' between language and use will guide me throughout the study.

Bybee (2006) proposes that in a usage-based theory "grammar is the cognitive organization of *one's experience with* language" , rather than organization of language (p. 711, emphasis added). It is through particular uses that the representation of language is built. Bybee (2006) argues for an exemplar representation of language experience in order to describe the type of cognitive representation of language, and notes that "constructions provide an appropriate vehicle for this

type of representation" (p. 712). The following paragraphs are devoted to a presentation of Exemplar Theory and its relevance to language.

As seen above, one of the offshoots of the usage-based model of language assumes an exemplar-based model of language (and cognitive) representation. This approach is mostly advocated in the works of Bybee (e.g., Bybee (2006); Bybee and Eddington (2006); Bybee (2010)). The exemplar-based model has been adopted in many linguistic fields, as evidenced in the 2006 special issue of The Linguistic Review (Gahl and Yu, 2006), devoted to exemplar-based models in linguistics. Thus, exemplar-based accounts are conquering more and more areas, such as language evolution (Wedel, 2006), syntactic acquisition (Abbot-Smith and Tomasello, 2006), or contrastive analysis of constructional spaces in different languages (Levshina, 2012; Levshina et al., 2013).

In an exemplar-model of language, exemplars are actual tokens of linguistic experience. These are categorized and matched with similar tokens of experience that have previously been stored as exemplars. In this way, the effect of token frequency comes about since high frequency exemplars are stronger than low frequency ones. The strength of the exemplar results in: (1) relative ease of access, and (2) increased relative morphological stability. Each exemplar of a morpho-syntactic construction includes information about the contexts of use and this includes the inferences made in such contexts.

Exemplars (instances of use) are categorized through parameters of frequency and similarity. An instance of use is inserted into the category and is mapped relatively to already existing exemplars (in terms of similarity). Exemplars showing high frequency can get grammaticized through a salient discourse pattern (Ariel, 2008). And moreover, constructions that are highly grammaticized become both highly schematic and productive: as an exemplar of a construction reaches high token frequency, it is processed without activating the other exemplars of the construction and begins to lose in analyseability and compositionality.

A potential obstacle facing the linguist when adopting an exemplar-based model of language in their analysis is the notion of similarity. Similarity in this context has to be quantified, or it cannot constitute a valid parameter of categorization. In recent years exemplar-based theories have been using quantitative methods to define similarity, for example utilising exploratory multivariate statistics on corpus data (e.g., Baayen, 2008). Such an approach is adopted in the present study as well. The following sections present the corpus serving as a database for analyzing the Hebrew Dative, and the quantitative method of analysis.

2.2 Data

The corpus serving as a database for the present research is an approximately 1,760,000 words corpus of spoken Hebrew. The corpus is a collection of transcriptions of 198 meetings of committees of the Knesset, the Israeli parliament, composed of multiple registers of language, both formal and colloquial. From this corpus, all occurrences of a Dative (*le-*) marked pronoun were extracted, yielding 16,575 tokens of Dative uses.[1] The process of downloading the files from the Knesset database and extracting only clauses containing the Dative was automated using an `R`-script I modified to fit my needs for web-crawling and text analysis (R Core Team (2014), following Baayen (2008); Gries (2009)). Each token was then coded by hand for 17 parameters (some were relevant only for a subset of the tokens). This is the list of parameters:

1. **Dative function**. 9 values were found in the corpus: ADD (Addressee), AFF (Affectee), DM (Discourse Marker), ETH (Ethical Dative), EVAL (Evaluative Reference Point), EXP (Experiencer), HEP (Human Endpoint), POSSDAT (Possessive Dative), REC (Recipient). See chapter 3 for a description of each function.

2. **Number of arguments**. 3 values were found: one, two, and three arguments.

3. **Word order**. 2 values were coded: predicate-first vs. subject-first

4. **O argument**. The lexical category of the O argument (if one occurs). Six values were found in the corpus: A (Adjectival), CL (finite clause), IRR (no O argument), NP (Noun Phrase), P (Prepositional Phrase), and V (non-finite clause).

[1] The choice of extracting and analyzing only pronominal Datives is motivated by several reasons: First, a substantive reason: the present thesis focuses on human Datives, leaving aside the discussion about adverbial Datives. A corpus search for pronominal *and* lexical Datives would yield both human referents and adverbial uses of the Dative, such as the *allative* or the *purposive*. Thus, I performed a search for pronominal Datives which can only refer to human referents. Second, a statistical reason: I conducted a pilot corpus search, randomly selecting 1,000 hits of all *le-* marked words, and found that 538 tokens were irrelevant for our purpose, encoding adverbial uses, as in *le-sham* 'to-there', or *le-maxar* 'to-tomorrow'. Among the relevant uses, encoding participant roles, 246 were pronouns and only 182 were lexical nouns. And moreover, the relative frequency of the various functions served by the Dative was found to be the same in both the pronominal and the lexical Datives. That is, a pronoun-focused search was needed for clearing the corpus from irrelevant hits, while according to the pilot search no essential data has been lost since the ratio of Dative functions within the lexical uses was found to be the same as in the pronominal uses.

5. **Agency of the A/S referent**. Four values were coded: IRR (irrelevant, for clauses with no A/S argument), low, medium, and high.

6. **Affirmation of the clause**. 2 values: Affirmative vs. Negative.

7. **Ellipsis**. For elliptic argument cases, in which three-argument clauses appear as two-argument clauses, and two-argument clauses appear as one-argument clauses. 2 values: ellip.no, and ellip.yes

8. **Mode of the clause**. Realis vs. Irrealis.

9. **Lexical category of the predicate**. Nine values were found in the corpus: TRANS.V (Transitive Verb), INTRANS.V (Intransitive Verb), ADJ (Adjective), ADV (Adverb), COMP.V (Complex Verb), D.M (Discourse Marker), INTERROGATIVE, N (Noun), and P (Preposition).

10. **Binyan**. Morphological paradigm of the predicate. Eight values were found: *Kal, Hifil, Piel, Pual, Hufal, Nifal, Hitpael*, and Irr, for non-verbal predicates.

11. **Voice of the predicate**. Passive vs. Active.

12. **Person of the Dative referent**. Three values: 1^{st}, 2^{nd}, and 3^{rd}.

13. **Direct Object marker on O argument**. Three values: Yes, No, and Irr, for clauses with no O argument.

14. **Individuation of the O referent**. A combination of parameters, see section 2.1.3. Five values were found: Irr, Low, Mid-low, Mid-high, and High.

15. **Construction**. Argument structure of the clause. 56 variants of structures were found.

16. **Verb and Root**. Each token was coded for its predicate, and the morphological root of the predicate, as there may appear two different verbs with the same morphological root in two different environments. 485 verbs were found in the corpus, derived from 343 morphological roots (see Appendix A for a complete list of the predicates in the corpus).

17. **Semantic type of predicate**. The verbal concept of the main predicate, along the lines presented in section 2.1.2. Eight types were found in the corpus: six verbal concepts (Primary A, Primary B, Secondary A, Secondary B, Secondary C, and Secondary D), and two adjectival types (Value and

Property). The semantic type of the other lexical categories (see item 9) was coded as Irr (see Appendix A for a full list of the predicates in the corpus according to their semantic types).

2.3 Method: Multiple Correspondence Analysis and Hierarchical Clustering on Principal Components

The usage-based approach to language is embedded in Cognitive Linguistics, a model of language that emphasizes the role of method and data (Glynn and Fischer, 2010, p. 2). Thus, a usage-based model of language calls for an empirical, quantitative approach to linguistic research, and specifically to corpus linguistics. Quantitative methods in linguistics aid in describing, explaining, and predicting linguistic phenomena (Gries, 2013). The analysis presented in this thesis is based on multivariate exploratory statistics, and specifically on Multiple Correspondence Analysis and Hierarchical Clustering on Principal Components.

Multiple Correspondence Analysis (MCA) is a technique for uncovering patterns in a multivariate database (Greenacre and Blasius, 2006; Abdi and Valentin, 2007; Husson et al., 2011). That is, it is an exploratory method used to detect patterns and correlations in data with multiple categorical variables. For example, collecting data such as answers to a questionnaire, the researcher can produce a table with columns representing the questions and rows representing individuals answering these questions. Each cell in the table, then, corresponds to an individual's answer to a particular question. In a linguistic corpus, the rows of such a table would be hits of the relevant search in the corpus. These are in fact the individuals of the data. Each column in the table would be a parameter each hit is coded for. These are the categorical variables of the data. The variables are categorical since they have several possible categorical, non-continuous, values. For example, the variable *number of arguments* in the current corpus is a categorical variable with three values: one-, two-, and three-arguments. Each cell in the table, then, indicates what value (category) a particular token possesses with regard to a particular variable. That is, in linguistic terms we can think of MCA as an "exploratory technique that reveals frequency-based associations in corpus data" (Glynn, 2012), visualizing these associations in the form of a map. Highly associated forms, for instance, appear closer on the map than forms with no association.

Studying the tokens in a corpus using MCA means finding the similarities

between sets of tokens that share variables. Thus, a new–statistically motivated–perspective on the corpus can be provided: similar tokens, and groups of similar tokens in terms of the categorical variable they share, can be found. Moreover, one can point to consistency in similarity, since MCA allows one to find what parameters determine similarity and what parameters are irrelevant for comparison between tokens. For example, two tokens that share each and every parameter will be considered highly similar:

(4) a. *holxim lishtot li et ha-maym me-ha-shorashim shel*
go to.drink to.me ACC the-water from-the-roots of
ha-cmaxim.
the-plants.
'They're gonna drink my plants' water (= they are going to draw water from a nearby underground location, whereby my plants will lose their water supply).'
(882, these numbers refer to the example number in the corpus)

b. *axarey ze baim ve-shotim lahem et ha-sibsud ha-ze*
after that (they).come and-drink to.them ACC the-subsidy the-this
im toxniot pratiot.
with plans private.
'And then somebody will waste their subsidy with private plans.
(2,562)

In (4) we see two sentences with the same main verb (*shata* 'drink'), a low individuated A referent, an O referent with a medium-high level of individuation, and an affirmative, irrealis clause. A Multiple Correspondence Analysis will treat these two tokens as highly similar, thus reflecting speakers' intuitions about their similarity.

Conversely, consider the following everyday usage of the Dative as found in the corpus:

(5) *shalom laxem.*
Hello to.you.
'Hello all.'
(3,680)

Intuitively, one can't argue that the sentences in (4) and (5) are similar on any level, apart from the fact that they all share a use of the Dative. And indeed, according to the many parameters they do not share, a Multiple Correspondence Analysis will treat these tokens as highly dissimilar.

In order to better understand the process of performing a Multiple Correspondence Analysis, let us assume that our corpus consists of only three tokens, the ones presented in (4)–(5). The coding table for this corpus is provided in Table 2.1. Looking at Table 2.1 we can see that the two similar tokens (882 and

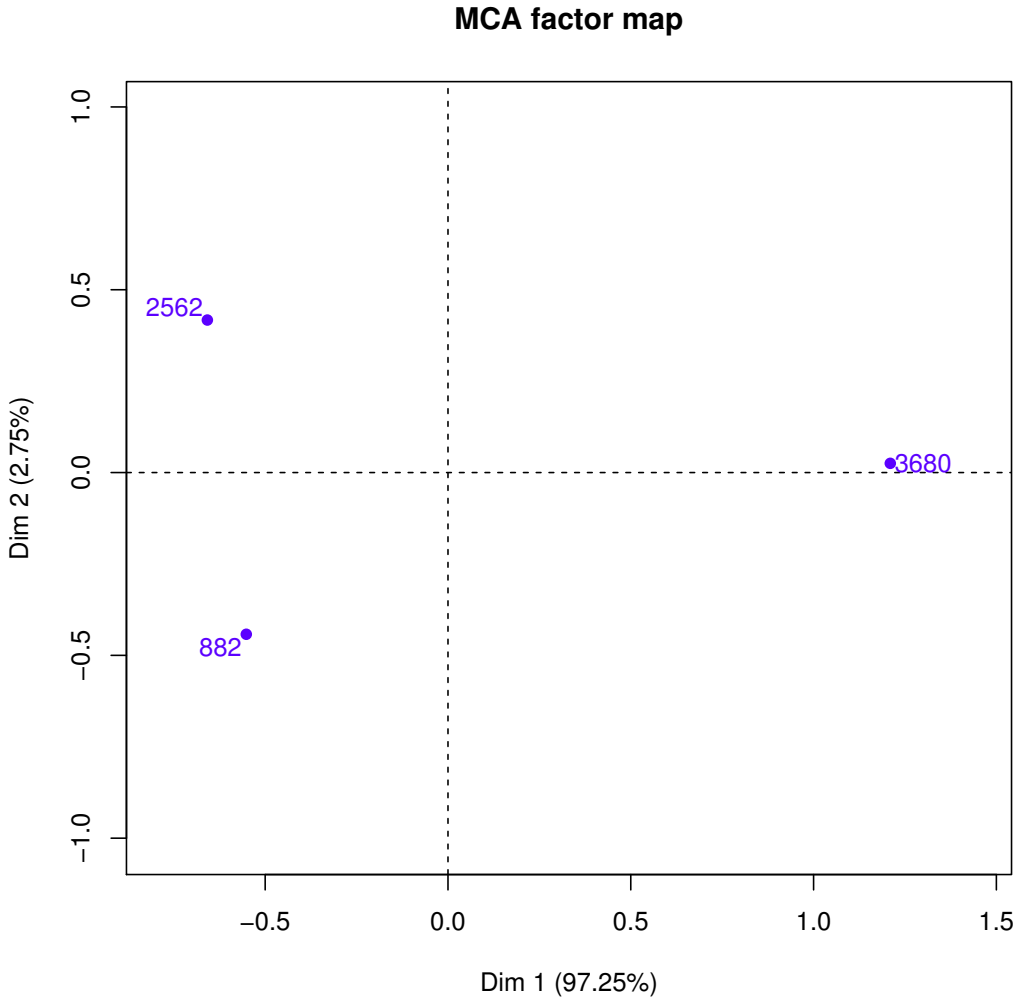
Table 2.1: A fragment of the dataset; coding for three tokens

ID	DAT.FUNC	AFFIRM	DAT.PRSN	ELLIPSIS	MODE
882	POSSDAT	affirmative	frst	ellip.no	irrealis
2562	POSSDAT	affirmative	thrd	ellip.no	realis
3680	DM	affirmative	scnd	ellip.no	irrealis
ID	NO.D	NO.OF.ARGUMENTS	O.DEF	O.INDIV	P.LEX.CAT
882	NP	three	O.def.yes	O.mid-high	TRANS.V
2562	NP	three	O.def.yes	O.mid-high	TRANS.V
3680	IRR	one	O.def.irr	O.irr	N
ID	P.TYPE	P.VOICE	PRED.FIRST	S.AGE.INDIV	V.BIN
882	pa	active	predNotFirst	S.low	KAL
2562	pa	active	predNotFirst	S.low	KAL
3680	IRR	active	predFirst	S.irr	IRR

2,562, which are only intuitively similar, so far) share almost every category, except for *mode* (realis vs. irrealis) and *dat.prsn* (person of the Dative referent). The third token, on the other hand (3,680), shares very few categories with the first two: *affirm* (whether the clause is affirmative or not), *ellip* (whether there is an elliptic argument or not), *mode* (with only one of the first two), and *voice* (active vs. passive). So far, then, just from looking at the coding table, we can conclude that the variables *affirm*, *ellip*, *mode*, and *voice* do not contribute to the differences (and similarities) between the tokens that were noted above, while the categories of other variables are strongly linked to either of the tokens in (4)-(5). Analysing a Multiple Correspondence Analysis on this dataset allows us (1) to represent these conclusions visually, and (2), to quantify the exact strength of a link between a category an an interpretation, the similarities between tokens, and the correspondences between categories.

First, consider the cloud of individuals in the data – a visual representation of the similarities between the tokens in the corpus (Figure 2.1). From the map in Figure 2.1 we can learn about the similarities and differences between the tokens. The map is composed of a horizontal and vertical axes, each marked with a certain percentage. These are the percentages of explained variance of the data each dimension provides (also termed 'inertia'). Thus, the horizontal dimension (dim

Figure 2.1: A Multiple Correspondence Analysis example with three tokens: correspondences between tokens



1) explains 97.25% of the variance in the data, while the vertical dimension (dim 2) explains only 2.75% of the variance. And indeed, we can see that the map represents the intuitive conclusion regarding the dissimilarity between 3,680 (5), on the one hand, and 882-2,562 (4) on the other. This dissimilarity is represented as distant points on the first (horizontal) dimension. Note that the minor difference we have detected between 882 and 2,652 ((4a)-(4b) above) is also represented, but on the vertical, less explanatory second dimension.

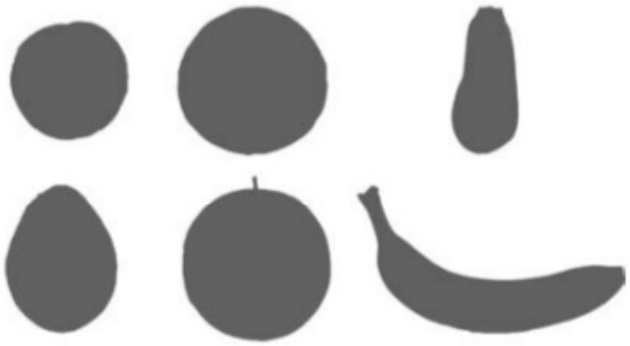
Note, however, that while tokens are compared on a basis of presence–absence of the categories, Multiple Correspondence Analysis takes into consideration the characteristics of every other token in the corpus when calculating distances. Thus, if two tokens share a rare category, they will be analysed as similar even though they present differences elsewhere in order to account for their common distinctiveness.

At this point of the method presentation let us consider the concept of dimensions and the variance they explain. In this context, I find Husson et al. (2011) explanation simple, clear, and illuminating. A variable is a dimension over which each token in the data is located on a particular point. Thus, each token in the corpus can be described according to each of the variables it is coded for. However, one cannot, and should not, describe each token according to each variable, since in most cases some variables are linked together, and some variables might turn out to be non-explanatory. Thus, one needs a tool that will find these links and produce the minimal number of dimensions required for a sufficient representation (i.e., explanation) of the data. Multiple Correspondence Analysis (among other techniques) is such a tool: it aims to study all the variables simultaneously, finding links between them. The solution provided by a Multiple Correspondence Analysis is usually a two-dimensional solution, providing the two dimensions that explain most of the variance in the data.

Husson et al. (2011) provide a clarifying, visual example for the selection of the dimensions that explain the most of the variance. Let us assume we need to represent a three-dimensional object, a fruit, in a two-dimensional manner, a sketch. One can either draw the object from the front, capturing its width and hight, or one can draw it from the profile, capturing its depth and hight. Each such drawing yields a different representation of the fruit. Consider Figure 2.2. The bottom row, the second two-dimensional drawing, is a better representation of the differences between the fruits. That is, it explains a greater percentage of the variance in the data. The same is true for corpus examples: most (in fact, almost all) of the variance in the small, three-token corpus is explained by the first dimension, reflecting the dissimilarity between 3,680 (5), and 882-2,562 (4).

Now that we have understood what the dimensions stand for, let us consider

Figure 2.2: Husson et al.'s (2011) Figure 1.5: Two-dimensional representations of fruits: from left to right an avocado, a melon and a banana, each row corresponds to a different representation.

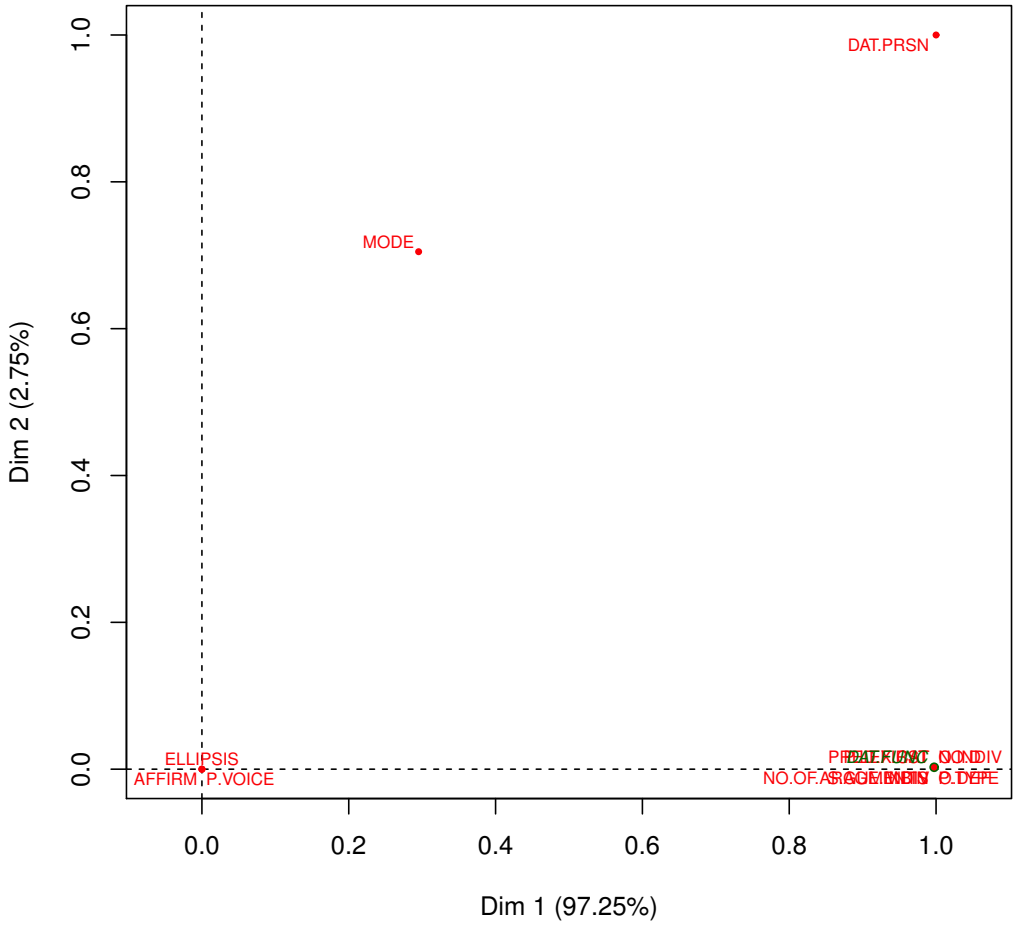


another plot provided by the Multiple Correspondence Analysis, representing the strength of the link between variables and dimensions, provided in Figure 2.3. A link between a variable and a dimension is represented as stronger as closer the variable is to 1 on the that dimension. Thus, the person of the Dative-marked participant, which is different in each of the three tokens, is strongly linked to both dimensions: it is located accordingly at the (1,1) point, differentiating each token from the others in this particular respect. The mode of the clause (realis vs. irrealis) differentiates between 882 and 2,562, but not between 882 and 3,680. Thus, it is linked mostly to the second, vertical dimension, located at (0.3,0.7). The other variables are either relevant only for the differences between the two similar tokens on one hand and the third token on the other, or irrelevant for any token. Thus, the variables *no.of.args* (number of arguments in the clause), *o.indiv* (individuation of the O participant), *p.lex.cat* (lexical category of the predicate), *p.type* (type of predicate), *pred.first* (linear order of Subject and predicate), *s.age.indiv* (Agency or individuation of the A/S participant), and *v.bin* (verb paradigm, *binyan*), are all located at the same point (1,0).

The irrelevant categories, which do not contribute to the interpretation of the tokens as similar or dissimilar, are located at the origin point (0,0). That is, since the first–horizontal–dimension explains almost all of the variance in the data, the categories that are linked to this dimension are the ones that contribute the most to the differentiation between tokens, and thus to their interpretation.

The third figure provided by the Multiple Correspondence Analysis represents the correspondence between the categories in the data (the different values of each variable), providing each category's average location (Figure 2.4). Apart from consideration of tokens, Multiple Correspondence Analysis allows one to analyse

Figure 2.3: A Multiple Correspondence Analysis example with three tokens: strength of link between variables and dimensions

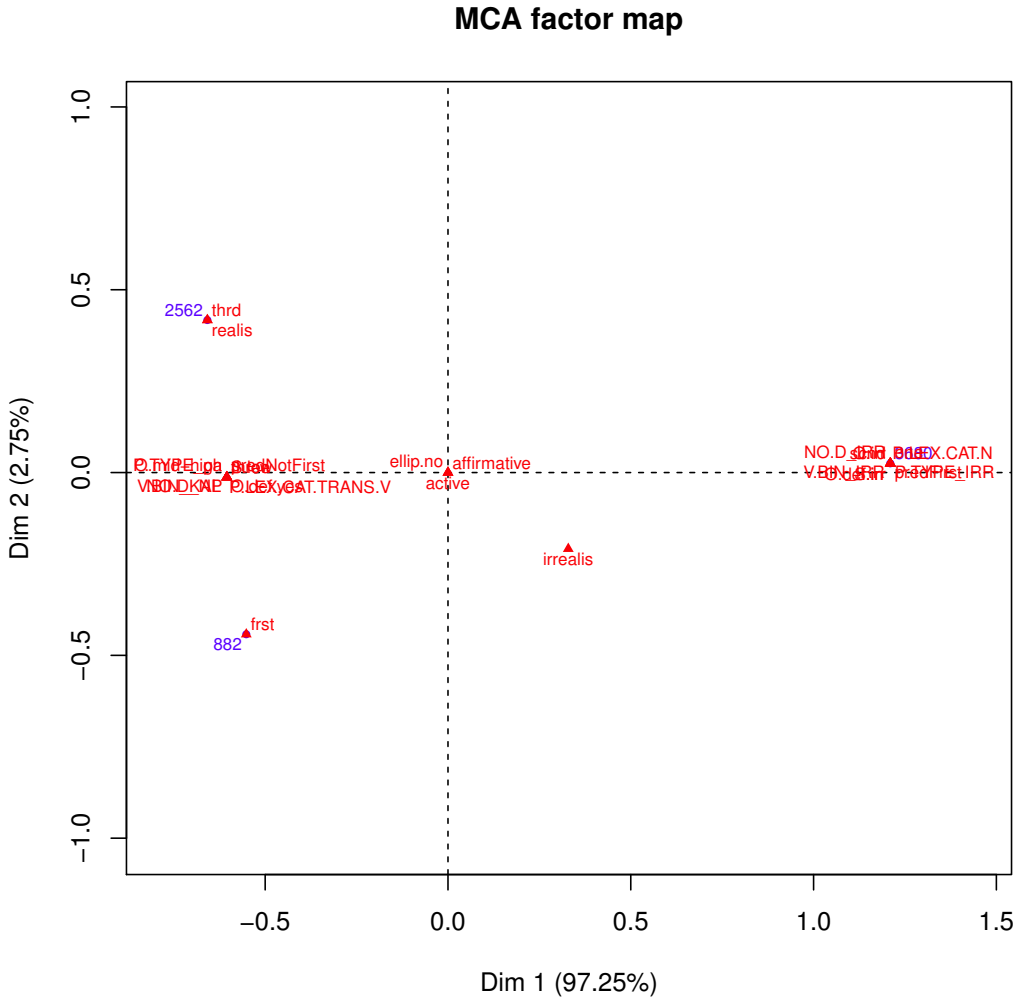


correspondences between categories. Each category represents both a variable and the set of individuals sharing this category. Two different categories (for example, a particular semantic type of predicate and number of arguments in the clause) appear adjacent to each other if they share the same set of tokens that select each category. The more individuals two categories have in common, the less distant they are from each other. Thus, while the category *irrealis* is located in the middle of the map, not corresponding with other categories, the categories *transitive verb, PA type of predicate, NP non-Dative argument, three-argument structure, definite O, medium-high individuation of O, predicate second, low agency of A/S,* and *KAL verb paradigm* highly correspond with each other, located at the same point on the map (-0.5,0). That is, these categories characterize the set of examples 882 and 2,562, located at the left half of the map. The other token in the corpus, 3,680, is mostly characterized by the categories one-argument structure, nominal predicate, neither S nor O arguments, and predicate first word order. These categories highly correspond with each other, located at the right half of the map.

Summing up this brief demonstration of the Multiple Correspondence Analysis process and outputs, we have seen that using Multiple Correspondence Analysis one can ask about correspondences in the data, and search for patterns. For instance, one can ask what is common between all tokens coded in the same way for the first, third, and fifth variables. Or, another possible question would be what is common between categories of different variables. That is, if, for instance, variable A has three possible categories (x, y, and z), and variable B has two possible categories (j, and k), are there any correspondences between the possible categories, such that A(x) and B(k) correspond with each other.

On a technical note, although MCA is an extension of Correspondence Analysis, distance between points is less meaningful than in correspondence analysis, due to the conflation of multiple dimensions into a two-dimensional map. However, quadrants and approximation do have meanings, and interpreting these plots is rather straightforward (Rencher, 2012; Glynn, 2012). In the present thesis I used the `R` program for statistical computing (R Core Team, 2014) with the package `FactoMineR` (Husson et al., 2013) for Multiple Correspondence Analysis and Hierarchical Clustering on Principal Components, with the MCA done on the Burt matrix. This package was chosen due to its visualization options and the possibility to add Hierarchical Classification on the basis of the MCA. While the explained inertia (explained variation in the data) is usually lower using this package than using the `ca` package (Nenadic and Greenacre, 2007) due to Greenacre's inertia adjustment, the number of variables in the present research renders the inertia almost un-interpretable (Glynn, 2012). Thus, better visualization options of the correlations and the possibility to compute the Hierarchical Classification on

Figure 2.4: A Multiple Correspondence Analysis example with three tokens: correspondences between categories



the MCA motivated my decision.

We have shown that using Multiple Correspondence Analysis one can visually observe that the data is divided into several clusters of tokens. However, such a visual observation is a subjective interpretation of the MCA map. In order to accurately describe the structure of the data and the grouping of tokens within it, one needs a statistical tool that partitions the data into sets of similar individuals. Hierarchical Clustering on Principal Components (HCPC) is such a tool. It is a complementary tool of Multiple Correspondence Analysis, used for graphically representing similarities or correlations (Husson et al., 2011), and for providing statistically defined clustering of tokens in the data, rather than the merely visual clustering possible through a Multiple Correspondence Analysis.

Hierarchical Clustering on Principal Components outputs a hierarchical tree in which similar observations sprout from the same branch. In the present thesis the hierarchical classification is done on the Principal Components. That is, it is based on the dimensions of the Multiple Correspondence Analysis that explain the structure of the data. The resulted clusters of tokens can be described by the variables in the data, the dimensions of the MCA, or the individual tokens that each cluster is composed of. That is, besides observing the visual output of the Hierarchical Clustering on Principal Components, one can interpret the numerical output of the clustering process, and learn which categorical variable is linked to which cluster.

For example, the variable *number of arguments* has three categories: one, two and three arguments in the clause. Hypothetically, using Hierarchical Clustering on Principal Components we might learn that the category *three arguments* is linked to cluster one, while the categories *one argument* and *two arguments* are linked to cluster two. This observation can teach us that the tokens with a three-argument syntactic structure share more features within themselves than they share with one- and two-argument tokens. That is, we can conclude that (i) beyond the fact that they share a particular category, they are similar (or dissimilar) on other levels as well, and (ii), even if two tokens do not share a category, they still belong to the same cluster (i.e., they are similar enough) since they share other features.

Thus, Hierarchical Clustering on Principal Components helps us find statistically defined clusters of exemplars based on the Multiple Correspondence Analysis. That is, using Hierarchical Clustering on Principal Components we can find a cluster, its prototypical exemplars, and the relevant features that provide its unique description. In addition to the graphical output, Hierarchical Clustering on Principal Components allows one to consider the most prototypical tokens in a cluster, such that they constitute the center of gravity: these are the tokens that are clos-

Figure 2.5: Hierarchical Clustering on Principal Components: tree diagram

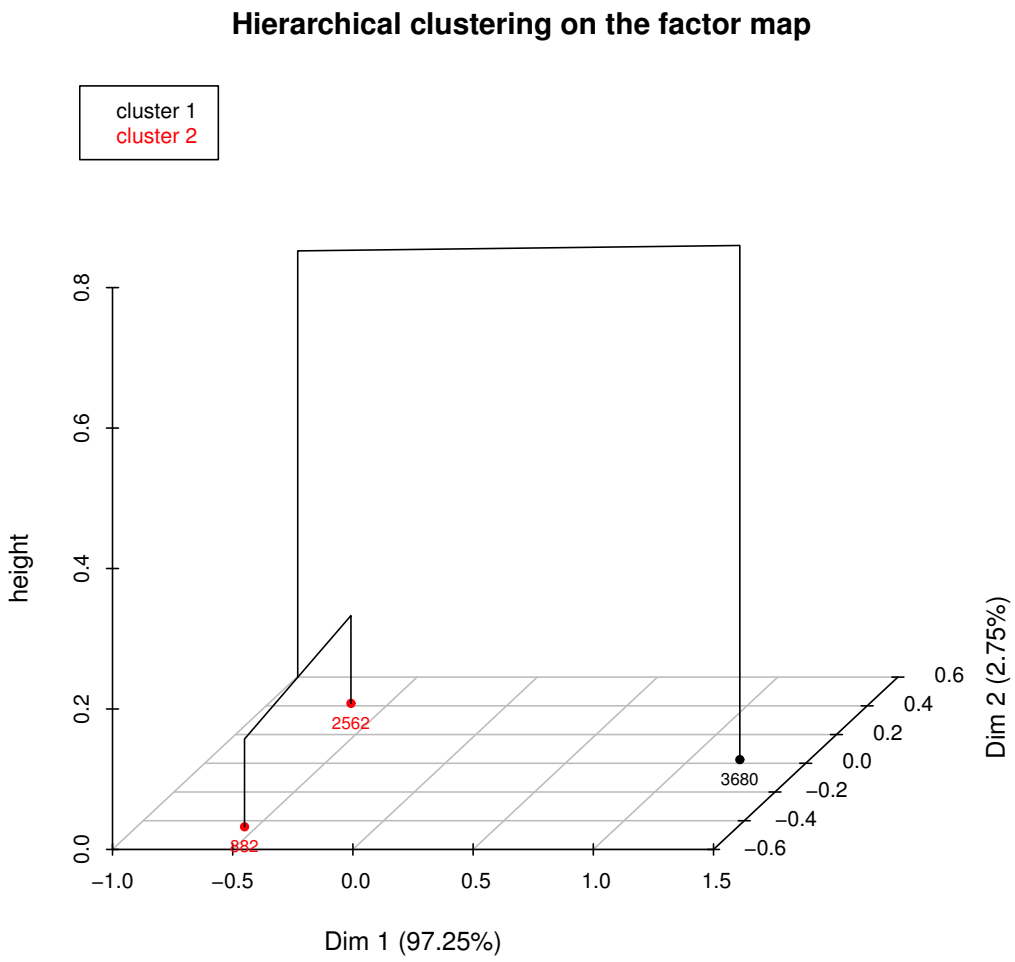
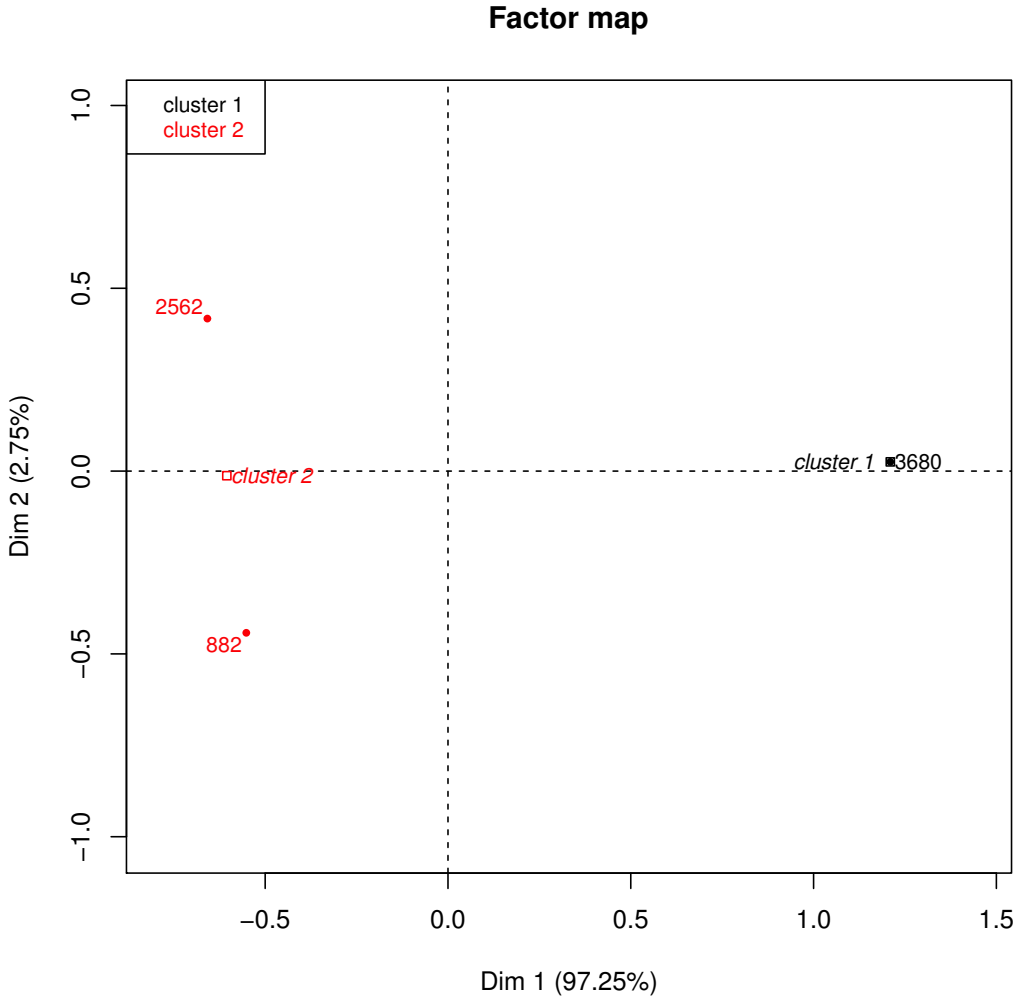


Figure 2.6: Hierarchical Clustering on Principal Components: two-dimensional map



est to the center of the cluster. All other tokens in the cluster are related to these tokens through a chain of family resemblances. Moreover, Hierarchical Clustering on Principal Components provides the objects that belong to a cluster and are placed the furthest from other clusters' centers. Since the pseudo-corpus used here to introduce the statistical techniques of Multiple Correspondence Analysis and Hierarchical Clustering on Principal Components consists of only three tokens, we cannot exemplify and explain the notion of central and unique exemplars in a cluster. This explanation is given as we perform Hierarchical Clustering on Principal Components on the actual corpus, in Chapter 5.

3
Dative participant roles in Hebrew

The present chapter surveys the participant roles marked by the Hebrew Dative as found in the corpus. As this study's main claim is that from a usage-based point of view we should look at discourse-based Dative functions, converging several similar participant roles together, the goal of the present chapter is to provide the basis for the multivariate corpus analysis section of the study, by spelling out the variety of participant roles the Dative can mark in Hebrew. Through the corpus analysis detailed in chapters 4 and 5 I argue that these participant roles can be merged into three sets of Dative functions, converged under three Discourse Profile Constructions.

In a seminal paper, Berman (1982) surveys the Dative in Hebrew, advocating an account of the Hebrew Dative as a marker of the Affectee of the event. Berman's analysis and list of Dative marked participant roles have been a starting point for studying the Hebrew Dative for many researches, the present one included.

Berman analyzes seven types of Dative uses, noting that the common semantic feature of these participant roles is marking the Affectee of an event: (i) Ethical Dative, (ii) Experiencer, (iii) Recipient, (iv) Benefactee, (v) Possessor, (vi) Deprivee, and (vii) Locatee. Berman notes that as long as an object argument can be perceived as affected by an event, such affectedness can be marked by the Dative. Berman emphasizes the discursive function of Hebrew Dative constructions, namely, as an alternative to the canonical constructions in Hebrew – both transi-

tive and intransitive – in which the Agent is downgraded (and see also Lyngfelt and Solstad (2006); Næss (2007)). Thus, the potential Agent is construed as an Experiencer of a situation, or an Affectee of an event, undergoing a change rather than performing one.

Berman (1982) distinguishes canonical Datives such as the Recipient in *give* or the Addressee in *tell* from extended Datives which are not required by the verb's lexical array – the Benefactee, the Possessor, the Deprivee, and the Locatee. The following are examples for Berman's extended Dative functions from the current corpus:

(1) a. *merkaz ha-mexkar ve-ha-meyda shel ha-kneset* ***hexin lanu***
The Knesset's center for research and information prepared to-us
niyar reka.
(a)paper background.
'The Knesset's center for research and information prepared a background paper for us.'
(2723) (Berman's Benefactee)

b. *efshar lenatek lo et ha-maym.*
possible to disconnect to-him ACC the-water.
'It's possible to cut off his water supply.'
(2180) (Berman's Possessor)

c. *ganvu lo maclema.*
stole to-him (a)camera.
'They stole a camera from him / They stole his camera.'
(9937) (Berman's Deprivee)

d. *ha-axot sama lo talk.*
the-nurse put to-him powder.
'The nurse put powder on him.'
Berman's (1982) Locatee, ex. (21-iv), not found in the present corpus.

Berman (1982) accounts for these roles as pertaining to the affectee of an event. I suggest to analyze them not necessarily as affected entities, but rather as *reference points given by the speaker in the process of construing the state of affairs*. In chapter 4 I present an affectedness scale over which Dative-marked participant roles (from different constructions) span, showing that a unified notion of Affectee is not enough; affectedness is a gradient concept of reference points for events: from a non affected entity only mentioned in order to anchor a situation to

a reference point, through partially affected entities that go through no (or minor) change of state, to a participant that goes through a change of mental state.

In order to reach this conclusion, however, we must first list the different functions served by the Dative in Hebrew as found and analyzed in the present corpus. In the following sections I go over the list of Dative-marked participant roles found in the corpus, exemplifying each role and discussing the differences between similar and related roles. This presentation of *Dative functions* is a necessary preliminary basis for the account of Hebrew *Discourse Profile Constructions* advocated in the present study.

Nine different Dative functions have been found in the corpus, with four occurring much more frequently than the others. Table 3.1 presents the frequencies of the Dative functions. Each function is defined, exemplified and explained in subsequent sections according to order of frequency.[2]

Table 3.1: Frequency of Dative functions in the corpus

Dative function	**Tokens**	**Relative frequency**
Addressee (ADD)	3,278	33.8%
Experiencer (EXP)	1,825	18.8%
Human Endpoint (HEP)	1,804	18.6%
Recipient (REC)	1,642	16.9%
Evaluative Reference Point (EVAL)	410	4.2%
Discourse Marker (DM)	303	3.1%
Possessive Dative (POSSDAT)	224	2.3%

Continued on next page

[2]While the complete list of Dative occurrences in the corpus consists of 16,575 tokens, the statistical results reported in the present research are based on a total of 9,694 tokens. The 6,681 tokens gap between the two totals is of only one type of Dative construction, composed of an existential predicate, a Dative marked pronoun, an a Noun Phrase, as in *yesh/ein/haya li sefer* 'I have/don't have/had a book' (lit. there is/there isn't/there was to.me a book).

These 6,681 existential Dative tokens were coded for each of the 17 parameters as well, but were excluded from the multivariate statistics due to two reasons. First, as a very large, homogeneous cluster, they present a behaviour that can be detected without the aid of the techniques used in the research. That is, their unique form, consisting of only three predicates, renders them non-competitors with any other Dative token as to their interpretation. Second, from a methodological point of view, the size of the cluster relative to other clusters in the corpus renders the other clusters too small to interpret. Moreover, Bybee (2010) shows that categories with such type-token ratio (of many tokens from a very small group of types) constitute a cognitive category on their own, separated from the original category they may have been part of (the Dative category, in the present case). Thus, while the results including the existential construction might have been reported, I have decided against it.

Table 3.1 – *Continued from previous page*

Dative function	**Tokens**	**Relative frequency**
Affectee (AFF)	159	1.6%
Ethical Dative (ETH)	49	0.5%
Total: 9 types	9,694	100%

3.1 Addressee

The Addressee role stands for the endpoint of a Telling or Showing event (Dixon, 2005; Baker et al., 1998). That is, it is the target towards which a message is delivered. In a sense, it is the *mental* endpoint of a Telling/Showing event (lexicalized by speaking and cognition verbs, respectively). Cross-linguistically, it is widely used in the ditransitive construction as an alternative to a Recipient (Haspelmath, 2011). In the present corpus, the Addressee (henceforth, ADD) is the most frequent Dative marked participant role with 3,278 tokens which are 33.81% of all Dative-marked participant roles. The following are representative examples for this function of the Hebrew Dative:

(2) a. *hi* ***amra li*** *she-hi yodaat.*
she told to-me that-she knows.
'She told me she knows.'
(10,220)

b. *ani roce* ***lefaret lexa ve-la-vaada*** *et ha-ceadim*
I want to detail to-you and-to-the-committee ACC the-steps
she-nakatnu me-az.
that-we took since-then.
'I would like to describe to you and to the committee each of the steps we've been taking.'
(12,376)

c. *harbe meod kalkelanim she-kol ha-zman* ***metaftefim lanu***
many economists that-all the-time drizzle to-us
she-asur lehaalot et netel ha-mas.
that-forbidden to raise ACC taxes
'Many economists that keep dripping on us that we shouldn't raise taxes.'
(12,346)

Example (2a) represents a prototypical use of the ADD role, construing a three-participant message delivering event. It is composed of a volitional, agentive, highly individuated Agent, a Speaking verb, and the content of the message as a clausal complement. The Dative marks a human entity, capable of receiving and comprehending a message. (2b) and (2c) illustrate extensions of the category to less prototypical cases of message-delivering. (2b) has a highly individuated A and O, but the Dative-marked referent is not a prototypical human entity. Rather, it is an institutional body, standing for and representing its members. The predicate is not prototypical for a three-participant construal as well; *lefaret*, 'to describe in detail', is mostly used in a two-participant environment. (2c) is an example for a novel use of the Addressee Dative, complementing an intransitive predicate. Used in this environment, the verb *tiftef*, 'drip, drizzle', gains a function of a message-delivering predicate with a speaker in the A syntactic relation, a message (and its content) in O position, and an Addressee: the recipient of the message, or the entity by which the message is intended to be comprehended.

3.2 Experiencer

The second most frequent Dative marked participant role is the Experiencer (henceforth, Exp), with 1,825 tokens (see Table 3.1). The Experiencer function is associated with cases in which the human referent is involved in a stative situation, not affected by an event. It indicates a person's relationship to the physical universe and how that person perceives or experience the universe. For instance, ((3a)) are two prototypical examples of the experiencer:

(3) a. *ha-yeda ha-ishi shelxa* ***nire li meod*** ***xashuv***.
the-knowledge the-personal yours seem.3SG.MS to.me very important.
'Your personal knowledge seems very important to me.'
(15,314)

b. ***koev li ha-lev*** *al ma she-kore b-a-maxon.*
hurts to.me the-heart on what that-happens in-the-institute.
'My heart aches with what's happening in the institution.'
(11)

The Experiencer occurs with a very wide range of semantic types of predicates. That is, contrary to other participant roles, the Experiencer seems to be very promiscuous as to what type of predicates can construe a state of affairs involving

such a participant. As such, this is a difficult role to give a clear-cut definition to. It functions as an umbrella under which a wide variety of participants can be grouped. Under the Experiencer umbrella we can find the experiencer of sensations, feelings, or happenings, the perceiver of impressions, and the cognizer of thoughts. Common to all Dative Experiencer participants is the feature of being related, not necessarily affected by, a state of affairs (and necessarily not an event). Looking at the predicates with which the Exp occurs in the corpus we can learn that it is mainly related to intransitive mental events, marked by adjectives or intransitive verbs. That is, using the Experiencer Dative the speaker profiles a human referent to which she construes a state of affairs as relevant, or to whom a certain feature is attributed. The Experiencer Dative can be employed in various semantic frames to mark a human endpoint relative to that frame. This particular characteristic of the Experiencer enables its use as an *added* endpoint, with almost no constraints.

The following is an examples for the Experiencer, emphasizing the *undergoer* nature of this participant role (Berman, 1982; Næss, 2007).

(4) *ze mafria lo meod.*
this bothers to.him very.
'It bothers him very much.'
(2,213)

The state of affairs in (4) can be conveyed in another way, using a transitive construal and different main predicate:

(5) *hu xoshev she-ze meod mafria.*
he thinks that-it very bothers.
'He thinks that this is bothering.'
(≈4)

The Dative-marked participant in (4) is the same referent as the A participant in the transitive paraphrase presented in (5). The substantial difference between the two examples lies in the construal of events paired with each construction. That is, while in the transitive construal (5) the referent is construed as an Experiencer of a mental action, a cognizer, with control over their thought, in the Dative construal (4) the referent is construed as having no control over the situation; they are experiencing a situation that exists on its own.

While the presentation of Dative functions in the present section follows their frequency order in the corpus, the following sub-section deviates from the order of frequency to present a participant role very similar to the Experiencer, namely,

the Evaluative Reference Point.

3.3 Evaluative Reference Point

The Evaluative Reference Point participant role (Eval) is one of the core functions of the Hebrew Dative; it marks a reference point against which a certain state of affairs is being evaluated. As such, it co-occurs mainly with adjectives and modal predicates, and has many features suggesting it is a subjective construction. Consider the following prototypical Eval examples:

(6) *lamadnu kama lekaxim* ***ve-asur lanu lishkoax otam***.
we learned few lessons **and-forbidden to.us to forget them**.
'We've learned some lessons that we shouldn't forget.'
(9,656)

(7) ***meod xashuv lanu she****-ha-universita o ha-mixlala tihye*
very important to.us that-the-university or the-college will be
mexuyevet l-a-inyan.
obligated to-the-issue.
'It's very important for us that the university or the college will be obligated to the issue.
(1,237)

One may ask what is the difference between (6) and (7) on one hand, and (3a) with the predicate *nire* 'seems' on the other. Seemingly, they all convey a perception of a situation. However, the Dative-marked participant in (3a) is in fact a cognizer whose thoughts are being reported, while the Dative-marked participant in (6)-(7) is a reference point for a modal reference of a situation. That is, an Experiencer Dative use does not necessarily trigger a modal reference (as in 6) or an evaluation of the situation (as in 7). Thus, (3a) can be paraphrased as:

(8) *ani xoshev she-ha-yeda ha-ishi shelxa meod xashuv.*
I think that-the-knowledge the-personal yours very important.
'I think your personal knowledge is very important.'
(≈(3a))

in which no evaluation is conveyed – only the content of the thought itself. (7), however, cannot be paraphrased without referring to the evaluation of the situation:

(9) *ani xoshev she-ze she-ha-universita o ha-mixlala tihye mexuyevet*
I think that-it that-the-university or the-college will.be obligated
l-a-inyan ze meod xashuv.
to-the-issue is very important .
'I think that the obligation of the university or the college is very important.'
(≈(7))

(9) shows us that the conveyed meaning, or the main message, is not the thinking event, but rather the evaluation. That is, the Dative referent in (6) or (7) cannot be interpreted as having some experience as a result of a situation, nor can it be said to be attributed a "forbidden" or an "important" quality. The Dative referent functions mainly as a reference point for the evaluation of a state of affairs – forgetting the lessons we've learned (6), or the importance of a situation (7). Had there not been a reference point, the assertion would have been an impersonal, objective declaration, as in:

(10) a. *lamadnu kama lekaxim ve-asur lishkoax otam.*
we learned few lessons and-forbidden to forget them.
'We've learned some lessons that shouldn't be forgotten.'

b. *meod xahuv she-ha-universita o ha-mixlala tihye*
very important that-the-university or the-college will be
mexuyevet l-a-inyan.
obligated to-the-issue.
'It's very important that the university or the college will be obligated to the issue.

Indeed, the same phenomenon can be found in the Experiencer Dative function:

(11) Experiencer:

a. *nimas li mi-kol ha-diburim.*
got tired to.me of-all the-talking.
'I got tired of all the talking.'
(6,643)

b. *nimas mi-kol ha-diburim.*
got tired of-all the-talking.
'All this talking is tiring.'

However, looking at the data we can see one main difference: while the Eval Dative co-occurs with modal predicates, the Experiencer deals with feelings, sensations, thoughts, or beliefs. That is, Evaluative Reference Point is a modal ref-

erence point; the source for the evaluation is external to the Dative referent. The source for the experience, the stimulus, may be external as well, but the experience itself (the feeling, sensation, etc.) is internal to the Dative referent. In the Evaluative Reference Point cases there is no internal experience; only an external evaluation of the state of affairs.

Summing up, while the Experiencer is a report of the Dative referent's attitude towards the world, how they experience or perceive it, the Evaluative Reference Point is an evaluation of a situation, not of the Dative referent itself. Using the Experiencer the speaker conveys what happens to the Dative referent, or how they experience a situation (for instance, as difficult or easy). With the Evaluative Reference Point interpretation, the utterance is about what the Dative referent thinks about a situation; how they evaluate a state of affairs: as being important, forbidden or allowed, as worthwhile, as fit, expensive, dangerous, or characteristic (see Table A in appendix A for a complete list of predicates occurring with the Evaluative Reference Point interpretation). Thus, the Experiencer and the Evaluative Reference Point represent two complementary perspectives on the world: the first about how we experience situations, and the second about how we evaluate state of affairs.

3.4 Human Endpoint

The Human Endpoint (HEP, see Langendonck 1998) is an endpoint of a two-participant event, marked as human. For example:

(12) *ein mi she-**yaarox lahem xatuna**.*
there.is.no who that-will.set to.them wedding.
'There's no one that'll marry them.'
(5,054)

In most cases it can be replaced by a Direct Object with a simple (more specific and not verb-object complex) verb, such as:

(13) *hu yexaten otam.*
he will.marry them.
'He'll marry them.'

In (12), there is no indication that the Dative-marked participant has a benefit (or a malefect, for that matter) as a result of the occurrence of the event. That is, (12) depicts a two-participant event with an actor and a human patient. The event itself is not a three-participant event. There are only two participants in the

objective event (hence the possible paraphrase in (13)). The construal, however, is such that the event is verbalized as a verb-object complex, thus rendering the argument structure a three argument one. In fact, in (12) we have an initiator instigating an action that is verbalized as a composed event using a complex verb: *arax xatuna* 'arranged a wedding' rather than *xiten* 'marry (causative)'. Hence, the NP in the Direct Object position is not a genuine O participant (Thompson and Hopper, 2001). The only real patient in the state of affairs is the Dative-marked participant. In this sense, it is a Dative Direct Object. A support for this account of (12) comes from comparing the same verb in other environments:

(14) *hu arax lahem et ha-shulxan.*
he set to.them ACC the-table.
'He set the table for them.'
(Constructed example)

In (14) the Dative-marked participant is indeed an Affectee, a benefactive. The difference between (12) and (14) lies in the nature of the Direct Object. *xatuna* 'wedding' in (12) is not a real patient of *arax* 'set'; it is not an individuated, concrete entity that goes through a change of state as a consequence of the Actor's actions. The table in (14), however, does go through such a change, and can be (and is) considered a fully affected patient of the verb *arax*'set.' Consider the following example:

(15) *hi gursha o she-hi* ***asta lo*** ***min orali****?*
she was expelled or that-she did to.him sex oral?
'Was she expelled or did she give him oral sex?
(6,899)

The dative in (15) is not a third participant; it is the only patient of a complex verb 'gave oral sex.' And consider the difference between the following:

(16) a. ***asu lanu*** *recax ofi.*
did to.us murder character.
'They blackened us.'
(9,184)
b. ***racxu*** ***lanu*** *et ha-ofi.*
murdered to.us ACC the-character.
'They blackened us. (Lit. They killed our character)'

The difference between (16a) and (16b) is similar to the one described for (12) and (14), but here we can observe another interesting fact: Using the same verb

in a higher transitivity context changes the interpretation of the Dative-marked participant. That is, using a simple (specific) rather than a complex verb, and marking the Direct Object as individuated (with the definite Accusative marker *et*) renders the Dative a Possessive Dative (16b); that is, the most suitable paraphrase for (16b) is:

(17) *racxu et ha-ofi shelanu.*
murdered ACC the-character ours.
'They blackened us. (Lit. They killed our character)'

Thus, the Human Endpoint is correlated with low transitivity while the Possessive Dative is correlated with higher transitivity. This conclusion is statistically backed-up later on in chapter 4.

Consider again the triad (12), (13) and (14). The HEP interpretation of the Dative-marked participant can be learned from paraphrasing the pseudo three-argument structure into a transitive two-argument structure. This paraphrasing is possible only when the Dative-marked participant is interpreted as a Human Endpoint – as an inherent part of the event; compare the following pairs:

(18) a. *hu asa lo bxina.*
he did to.him test.
'He tested him.'
b. *hu baxan oto.*
he tested him.
'He tested him.'

(19) a. *ani holex ve-ose la ikur.*
I go and-do to.it sterilization.
'I'm sterilizing it.'
(3,366)
b. *ani meaker ota.*
I sterilizing it.
'I'm sterilizing it.'

(20) a. *yihye mitkan hafaka shel maym ve-**yaasu** **lahem** **hatpala**.*
there will be facility production of water and-will-do to.them desalination.
'There's going to be a water production facility and the water will be desalinated.'
(4,154)

b. *yihye mitkan hafaka shel maym ve-**yatpilu***
there will be facility production of water and-will-desalinate
otam.
them.
'There's going to be a water production facility and the water will be desalinated.'

However, when the Dative-marked participant is a true third participant interpreted as an Affectee (that is, as being affected as a result of the occurrence of a completed event), this paraphrasing is impossible:

(21) Affectee

a. *hu **arax** **lahem** et ha-shulxan.*
he arranged to.them the table.
'He set the table for them.'

b. **hu **shilxen** otam.*
*he tabled them.
'He set the table for them.'

c. *ovdim she-**yaasu** **la** et ha-avodot ha-pshutot.*
workers that-will do to.her the-works the-simple.
'Workers that will do simple jobs for her.'
(5,852)

d. No paraphrase

e. ***asinu** **la** mesibat sium kshe-hi azva.*
we did to.her party finish when-she left.
'We threw her a farewell party when she left.
(8,871)

f. No paraphrase

Looking at the corpus, we see that the HEP Dative function is related to clauses in which no affected participant is profiled other than the Dative-marked participant. For example, while in (22) the speaker refers to a concrete object that will go through a change of state (or, more specifically, a creation), (23) profiles no such object. The Direct Object in (23) (*nezek gadol* 'great damage') is an abstract situation and the only affected entity is the Dative-marked participant. Thus, the Dative is interpreted as an Affectee in (22), and as a Human Endpoint in (23):

(22) *lokeax lanu zman lehitargen im* ***toxna xaxama she-taase***
takes to.us time to get organize with **software smart that-will.do**
lanu et ha-xituxim ha-ele.
to.us these cuts.
'It takes us some time to get some smart software that will cut these cuts for us.'
(5,660)

(23) *hesbarti lama* ***xok idud hashkaot hon ose lanu***
I.explained why **low encouragement investments capital does to.us**
nezek gadol.
damage great.
'I already explained why the capital investment encouragement low causes us a great damage.'
(5,655)

We can conclude, then, that the main substantial difference between HEP and Aff is the directness of the effect. In HEP the Dative-marked participant is affected directly as a consequence of the Agent's action, and forms an inherent part of the state of affairs. In the Aff cases, on the other hand, the Dative-marked participant is affected as a result of the Agent's action *on another entity*, and is part of the speaker's construal of events. Notice, however, that directness of effect does not equal completeness of effect.

A true transitive event includes an object that goes through a complete change of state. These events are lexicalized by core transitive verbs (Levin, 1999). For example, these verbs can be passivized:

(24) a. *hu hika nimracot et ha-yeled.*
he beat fiercely ACC the-child.
'He fiercely beat the child.'
b. *ha-yeled huka nimracot.*
the-child was.beaten fiercely.
'The child was fiercely beaten.'

This type of events can be attributed as a characteristics as well, describing a referent:

(25) hu yeled muke.
he child beaten.
'He is an abused child.'

Verbs such as *hirbic* 'hit' are non-core transitive verbs. Their object does not go through a complete change of state, thus they cannot be passivized, and their object cannot be characterized by means of an adjectival form of the verb:

(26) a. *hu hirbic l-a-yeled.*
he hit to-the-child.
'He hit the child.'
b. **ha-yeled hurbac.*
the-child was.hit.
'The child was hit.'
c. **ze yeled murbac.*
this child hit.
'This is a hit child.'

It is exactly this quality that is unique to human patients: they usually do not go through a complete change of state. That is, they *can* go through a complete change of state, but when one refers to events with two human participants it is usually done with a non-canonical marking of the direct object (Haspelmath, 2001).

The Dative marks that there is a subject (a psychological one, not a grammatical one) at the endpoint of the action, rather than an object. That is, when a patient is marked by a canonical grammatical object (with an Accusative marker, for instance), it is marked as a psychological object, not attributed with any human features. When a patient is marked by a non-canonical object marker, the Dative in our case, it is marked as a psychological subject. This mapping from the conceptual-psychological level to the grammatical level is responsible for the findings in Levin (1999) and Haspelmath (2001), and for the Dative object of *hirbic* 'hit' in our case, as well as *heemin* 'believe,' and *azar* 'help.'

To sum up, the verb's argument structure, and the differences between *hirbic* 'hit' and its synonym *hika* 'hit' suggest that this verb construes a situation in which the patient is human, or marked as possessing human features (as in a parent telling his child to 'hit the table' if he fell on it and got bruised).

3.5 Recipient

The Recipient participant role is only the fourth most frequent Dative-marked participant role in the corpus. It is straightforwardly defined as a participant that receives an object – either a physical or a metaphorical one. It is usually the second human participant in a Transfer event, together with the Donor and the

transferred Theme which changes location (and possession) from the Donor's to the Recipient's (Malchukov et al., 2010).

Considering the predicates occurring with the Recipient, it has a very low type-token ratio with only 45 types of predicates spreading over 1,642 tokens (2.7% TTR). That is, the Recipient participant role is very faithful and reserved to very specific scenarios, namely, Transfer events. Consider the following prototypical examples:

(27) Transfer of a concrete object:

a. *hu natan li teudat zehut.*
he gave to.me (an)I.D. card.
'He gave me an I.D. card.'
(16,760)

b. *ani dibarti im mishehi she-**natna li et ha-telefon**.*
I talked with someone that-gave to.me OM the-phone.
'I talked to someone that gave me the phone number.'
(16,091)

(28) Transfer of an abstract object:

a. *misrad ha-pnim natan lanu gibuy mishpati.*
Ministry of the Interior gave to.us backup legal.
'The Ministry of the Interior gave us legal backup.'
(16,294)

b. *tnu lanu et ha-chans le-maan hashem.*
give to.us OM the-chance for God's sake.
'Give us a chance, for God's sake!'
(16,513)

As can be seen in (28), the Recipient need not actually receive a physical object, and this participant role can be applied to a metaphorical transfer or to a non-willing Recipient (and see Goldberg 1995:146–147 for a discussion about the role of willingness as a semantic constraint on the Recipient in English):

(29) *anaxnu notnim lo trufot neged recono [...] va-afilu*
we give to.him drugs against his.will [...] and-even
niz'ey-xashmal notnim lo neged recono.
electric-shocks give to.him against his.will.
'We're giving him drugs, and even electric shocks against his will.'
(10,015)

While it is a core participant of Transfer-related verbs, the Recipient also oc-

curs in usage patterns in which the predicate is not by itself a transfer-related verb, yet it is accompanied by a Dative marked Recipient. That is, a Dative construal of a specific type of predicates involves profiling an entity as receiving an object, which in turn is construed as created or moved. There may be an intended Recipient for this object on the part of the Actor, but it is not necessary. Consider the following examples:

(30) a. *hem gam ken hosifu lanu kesef.*
they too added to.us money.
'They gave us money too (= added an amount on top of another amount).'
(5,339)

b. *ani ash'ir lexa po otek.*
I (will)leave for.you here (a)copy.
'I'll leave a copy here for you.
(13,310)

In none of these examples can we say that there is a Recipient as part of the core participants of the situation. Neither *hosif* 'add' nor *hish'ir* 'leave' are Transfer verbs. Rather, these are Transitive-motion verbs: two-participant verbs with a profiled Initiator as the Mover and an Endpoint as the Theme. Consider a Dative-less versions of (30):

(31) a. *hem gam ken hosifu kesef.*
they too added money.
'They added money too.'
(∼5,339)

b. *ani ash'ir po otek.*
I (will)leave here (a)copy.
'I'll leave a copy here.
(∼13,310)

The core situations remain the same in both (30) and (31); the only difference is the addition of a human entity construed as being affected as a consequence of the occurrences in the core situation. The particular effect in each case is dictated by the nature of the event; a human entity can be affected by a Transitive-motion event by receiving the moved Theme. We can see, then, that a Dative construal of a Transitive-motion situation involves profiling a Recipient, marking it with the Dative. That is, it is a non lexically-determined Recipient, strengthening its analysis as an Extension to Core.

At this point it is important to consider the relations between Transitive-motion

events and Transfer events. Both consist of an Initiator causing the motion of a physical object from point *a* to point *b*. The difference lies in the typical construal of the situation. Transfer events are lexicalized by three-participant verbs, thus profiling the goal of the motion – the Recipient (in Hebrew and in many other languages, see for example Newman, 1998; Croft et al., 2001; Francez, 2006; Levin, 2008; Rappaport Hovav and Levin, 2008). Transitive-motion events, on the other hand, may be lexicalized by two-participant verbs with no goal profiled. For example:

(32) *mi heziz et ha-gvina sheli?*
who moved OM the-cheese mine?
'Who moved my cheese?'

In (32) we can see a core transitive, two-place construction used to construe a Transitive motion event. Such a construction cannot profile the goal, the Endpoint, with an argument. Thus the Dative construal comes in handy, allowing speakers to profile the goal of the motion in a Transitive motion event.

3.6 Discourse Marker

The Discourse Marker Dative-marked participant cannot be termed a participant role. However, it occurs in the corpus, more often than the Possessive Dative, the Affectee, or the Ethical Dative. Thus, although it is not a participant role per se, it deserves a discussion. Consider the following examples:

(33) a. *taamin li she-ze nadir.*
believe to.me that-this rare.
'Trust me, it is rare.'
(892)

b. *ha-vaada crixa* ***nidme li*** *be-yuni lehacig et hemshex*
the-committee need **seems to.me** in-June to present ACC rest.of
ha-toxnit.
the-plan.
'The committee has to present the rest of the plan in June, **I think**.'
(5,606)

c. *toda raba lexa.*
thanks very much to.you.
'Thank you very much.'
(684)

d. *slax li, ata xay be-seret.*
forgive to.me, you live in-movie.
'Excuse me, but you don't know what you're talking about.'
(15,803)

The Discourse Marker Dative (DM) is used as part of a parenthetical comment conveying either the speaker's attitude towards their own utterance ((33a),(33b)), or a politeness marker ((33c), (33d)). Both types of discourse markers served by the Dative in Hebrew resemble another Dative function found in the corpus, namely, the Human Endpoint. However, the DM is part of the discursive level of the sentence rather than its semantic interpretation. That is, it concerns the speaker and hearer, rather than the participants of the construed event. In this respect, this is a subjectivity marker: it puts the speaker (or the hearer) onstage (Verhagen, 2005; Langacker, 2009). Indeed, 98% of the tokens interpreted as Discourse Markers consists of a first or second person Dative referent. In section 3.8 we will see that the Ethical Dative is very similar to the DM to the extent that they both mark some kind of subjectivity. Moreover, the multivariate analysis in chapter 4 reveals the interesting position both the Discourse Marker Dative and the Ethical Dative hold in the Dative category in terms of transitivity and subjectivity.

3.7 Affectee and Possessive Dative

Zúñiga and Kittilä (2010) discuss the Benefactive (termed Beneficiary) as an Affectee representative, and define it as follows:

> The beneficiary is a participant that is advantageously affected by an event without being its obligatory participant (either agent or primary target, i.e. patient). Since normally only animate participants are capable of making use of the benefit bestowed upon them, beneficiaries are typically animate (p. 2).

Smith (2010) defines the Benefactive *constructions* in a similar way:

> [B]enefactive constructions explicitly indicate an event as good for a participant using some kind of morphosyntactic means (p. 73).

The Hebrew Dative Affectee fits these definitions, reflecting a generalized extension of an affective situation.

An affective situation prototypically involves two participants, the Initiator and the affected Endpoint. We have already discussed a Dative construal that allows the speaker to profile an extra participant, namely, the Recipient of Transitive

Motion events. Another type of participant which is relevant in this respect is the Affectee (34):

(34) *ein madad exad she-yaxol* ***lesakem lexa et ha-tifkud***
there is no estimator one that-can **summarize to-you the function**
shel ha-maarext.
of the-system.
'There's no one estimator that would **summarize the system's function for you**.'
(13,874)

The core situation in (34) is of summarizing: an Initiator creates an abstract object – the affected Endpoint. In such an event there is no intended Endpoint on the part of the Actor apart from the summary itself. This is a two-pole situation; a clear case of a two-participant event. However, we can see that the speaker in (34) profiles another participant, thus anchoring the situation to a third pole. The third pole functions as a reference point assigned by the speaker for the audience for interpreting the situation. That is, the audience is to interpret the event of summarizing with respect to the Dative referent. A situation in which there is a created or moved object can be interpreted relative to a third, extra, participant by construing the creation or motion as affecting the extra participant. In the case of (34), the summary is construed as being done *for the Dative referent to learn from*. This is the Hebrew Dative Affectee function.

The following example demonstrates the fine line between the Recipient participant role discussed above and the topic of the present section, the Affectee:

(35) *ha-medina* ***makciva lo dira,*** *moseret lo dira*
the-state **budgets to-him apartment**, delivers to-him apartment
be-haskara.
in-rent.
'The state budgets an apartment **for him**, provides him with a rental apartment.'
(2,515)

The English translation of (35) involves a *for* benefactive, which in Hebrew is marked by the Dative just as the Recipient is. However, the final clause of the sentence is a paraphrase of the first clause, where the non-Transfer verb (budget) is replaced by a prototypical Transfer one – *masar* 'deliver'. That is, it seems that the Affectee and the Recipient are almost interchangeable.

The Possessive Dative is a well studied Dative function, both cross linguistically and in Hebrew (e.g. Borer and Grodzinsky, 1986; Van Belle and Langendonck, 1996; Šaríc, 2002; Cuervo, 2003; Linzen, 2009; Gafter, 2014). Contrary to the major role it plays in the literature, the Possessive Dative occurs only 218 times in the present corpus (2.25% of all Dative clauses). The Possessive Dative refers to a human entity Y that is affected by an action performed by an Actor (X) upon an object (Z) *that belongs to Y*. The possession relations is the core difference between the Possessive Dative and the Affectee. Consider the following example:

(36) *soxney mishtara mazminim na'arat livuy ve-az* ***bodkim la*** ***et*** ***ha-niyarot***.
police agents invite escort girl and-then **check** **to-her** **ACC** **the-papers**.
'Police agents invite an escort girl and **check her papers**.'
(2,374)

The speaker in (36) construes the 'checking papers' event relative to the papers' owner, which is not a core participant of the situation. That is, no core meaning would be lost had the speaker uttered the Dative-less (37). Moreover, since the papers' owner is already mentioned, one might say that the possession relation remains as well:

(37) *soxney mishtara mazminim na'arat livuy ve-az* ***bodkim et*** ***ha-niyarot***.
police agents invite escort girl and-then **check** **ACC** **the-papers**.
'Police agents invite an escort girl and **check papers**.'
($\sim$ 2,374)

This is, then, another case of an extension to core. The dative construal allows the speaker in this case to profile the owner of the object of an affective action, thus conveying that the completed (core) action affected the owner of the action's Endpoint.

An important question should be raised at this point. Is it possession relations that the speaker conveys, or affectedness? The interesting debate regarding the Possessive Dative's definition and the importance of possession in this respect is crucial for the understanding of this Dative function, and for differentiating it from other functions such as the Affectee. Linzen (2009) deals with this exact issue, concluding that Affectedness is the main feature of the so called Possessive

Dative. I concur with this conclusion. Indeed, it is a main claim of the present dissertation that particular sets of Hebrew Dative functions cannot be categorically differentiated. In the present research Dative functions are accounted for within a larger context of a network of functions with the aim of showing that it is *discourse profiles* that are essential for the interpretation of the Dative. These patterns subsume several Dative functions converged together, such as the Possessive Dative and the Affectee.

For now, then, the present description of the Possessive Dative is adequate. I will go back to the nature of the relations between the Affectee and the Possessive Dative in chapter 4.

3.8 Ethical Dative

The Ethical Dative (Eth) is another Dative function that draws much attention due to its special status as a non-participant, subjectively portrayed as related, or indirectly affected by the event. Like the Possessive Dative, the Eth too is very rare in the present corpus: only 49 tokens of all the clauses containing a Dative-marked participant are interpreted as having an Ethical Dative function (0.51% of all dative tokens). Such a low frequency function results in a very small, non-structured category. Consider the following examples:

(38) a. *hu mevi lanu netunim* ***eize talmid nafal lo***
he brings to.us data **which student dropped out to.him**
beyn kita tet le-kita yud.
between ninth grade and tenth grade.
'He gives us data about students that dropped out on him between the ninth and the tenth grades.'
(178)

b. *eyze metayel yirce lehagia le-shetax tiyuley ofnaym*
what traveller will want to arrive to-field trips.of bicycle
kshe-kodxim lo, *hu xoshesh mi-zihum svivati.*
when-drill to.him, he afraid of-pollution environmental.
'What traveller would like to have a bicycle trip in a drilling area? He's afraid it would be polluted.'
(1,457)

Berman (1982) defines the Ethical Dative in Hebrew as an "outsider or mere onlooker, a nonparticipant who is nonetheless perceived as affected by the event described." It resembles the Discourse Marker Dative presented in section 3.6 in

that they both put a nonparticipant, usually one of the interlocutors, onstage (92% first and second person referents in the Ethical Dative case), thus subjectively construing a state of affairs. The difference between these two functions is that the referent of the Ethical Dative is construed as affected by the state of affairs, while the referent of the Discourse Marker is construed as evaluating the utterance itself. In other words, while the Ethical Dative is an outsider or an onlooker of the event (Berman, 1982), the Discourse Marker is an onlooker on the speaking event itself.[3]

3.9 Summary

The present chapter provided a list of Dative participant roles as found in the corpus. We have seen that in some cases it is difficult to tell apart one interpretation from the other, for instance the differences between the Affectee and the Possessive Dative, or between the Experiencer and the Evaluative reference point. While these fine-grained semantic differences between closely related participant roles is a topic for research in many cases, the current study suggests that from a usage-based point of view these different roles seem to be sub-varieties of a single discourse pattern. This conclusion is made based on the corpus analysis presented in the following chapters.

[3]In a recent account of Hebrew non-obligatory Datives, Boneh and Bar-Asher Siegal (2014) add an interesting characterization of the Hebrew Ethical Dative, noting that it is associated with an unexpected information.

4
Corpus analysis

In the previous chapter we have investigated Hebrew Dative functions from the semantic perspective of participant roles. It was shown that in many cases it is difficult to decide which participant role the Dative marks, and asked whether this decision is relevant from a usage point of view. In the present chapter we open the window into the Hebrew Dative in two steps. First, in section 4.1 we ask whether an approach concerning Argument Structure Constructions (Goldberg, 1995) is adequate for accounting for the Hebrew Dative's distribution and interpretations in the corpus. Then, in section 4.2, after showing the shortcomings of this approach we open the window even wider, considering discursive parameters, claiming that multifactorial discourse patterns are the right perspective to account for the Hebrew Dative usage patterns.

4.1 From Dative functions to Argument Structure Constructions

Figure 4.1 on page 57 presents a Cohen Friendly Association plot (Friendly, 1992) describing which Dative interpretation is associated with which syntactic structure, with a proportional presentation of the deviations from the expected frequency. Blue, rising bars, indicate a higher-than-expected association between Dative interpretation and syntactic structure, while purple, descending bars, indi-

cate a lower-than-expected association. A summary of these associations together with a reader for construction numbers is given in Table 4.1, followed by a list of examples for each syntactic structure.

Looking at the distribution of blue bars we can see that among the fourteen most frequent syntactic structures in the corpus (covering 95% of the tokens), most Dative interpretations are associated with one or two syntactic structures.[4] For instance, the Addressee is associated with structure 42, the Affectee with structure 43, and the Ethical Dative is associated with structure 32.

Table 4.1: Summary of associations in Figure 4.1

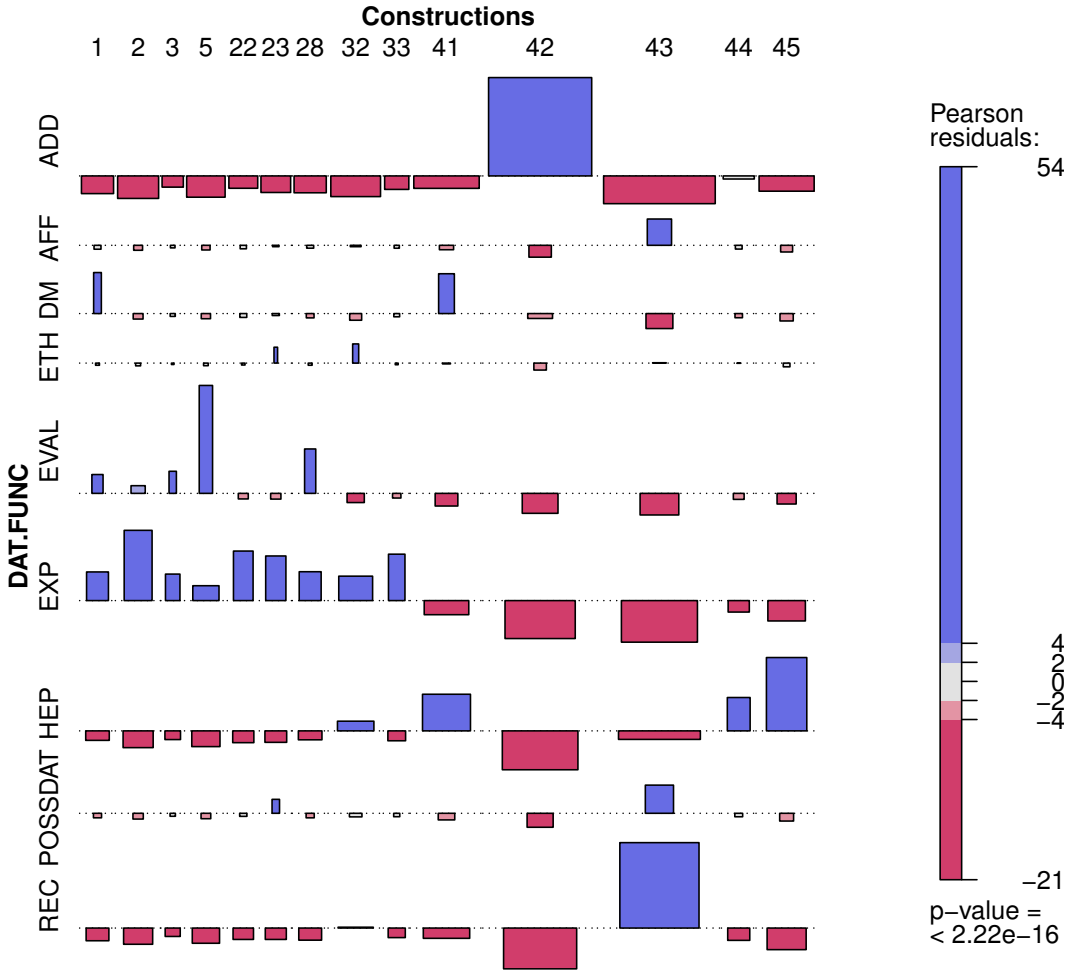
Number	**Syntactic structure**	**Dative function**
1	ADJ DatNP	Exp, Eval, D.M
2	ADJ DatNP CL	Exp
3	ADJ DatNP NP	Exp, Eval
5	ADJ DatNP V	Eval, Exp
22	INTRANS.V DatNP CL	Exp
23	INTRANS.V DatNP NP	Exp, PossDat
28	NP ADJ DatNP	Eval, Exp
32	NP INTRANS.V DatNP	Eth, Exp, Hep
33	NP INTRANS.V DatNP A	Exp
41	NP TRANS.V DatNP	D.M, Hep
42	NP TRANS.V DatNP CL	Add
43	NP TRANS.V DatNP NP	Rec, Aff, PossDat
44	NP TRANS.V DatNP P	Hep
45	NP TRANS.V DatNP V	Hep

Syntactic structure: ADJ DatNP (structure 1)

(1) *ha-yazam* *kamuvan* *yirce* *la'asot* *et* *ze* *eyfo*
the-entrepreneur obviously will.want to do ACC this where
*she-**noax*** ***ve-kal*** ***lo***.
that-**convenient and-easy for.him**.

[4] These fourteen syntactic structures constitute 23% of all types of structures found in the corpus, sixty in number. That is, 23% of the types cover 95% of the tokens. Moreover, the rest of the 5% of the tokens are mostly word order variants on one of the fourteen structures presented in Table 4.1 and Figure 4.1.

Figure 4.1: Cohen-Friendly Association plot: Dative function and syntactic structure



Reader for Dative functions (DAT.FUNC): Add: Addressee, Aff: Affectee, DM: Discourse Marker, ETH: Ethical Dative, EVAL: Evaluative Reference Point, EXP: Experiencer, HEP: Human Endpoint, POSSDAT: Possessive Dative, REC: Recipient

'The entrepreneur will probably want to do it wherever he feels most convenient.'
(2,771)

Syntactic structure: ADJ DatNP CL (structure 2)

(2) ***nidme li*** ***she****-shney yishuvim aravyim zaxu.*
seems to.me that-two towns Arab won.
'I think two Arab towns won.'
(156)

Syntactic structure: ADJ DatNP NP (structure 3)

(3) *lo* ***shave li*** ***ha****-od xodshaym ximoterapya.*
not worth to.me the-more two months chemotherapy.
'Two more months of chemotherapy are not worth it to me.'
(525)

Syntactic structure: ADJ DatNP V (structure 5)

(4) ***xashuv lanu lehagia*** *livney ha-noar.*
important for.us to reach to.sons of-youth.
'It's important for us to reach teenagers.'
(4,580)

Syntactic structure: INTRANS.V DatNP CL (structure 22)

(5) ***nire li*** ***she****-ze ha-davar ha-codek beyoter la'asoto.*
seems to.me that-it the-thing the-right modt to do.
'It seems to me that this is the right thing to do.'
(542)

Syntactic structure: INTRANS.V DatNP NP (structure 23)

(6) ***nishar lanu shetax*** *nifla be-ezor shfelat yehuda.*
left to.us territory wonderful in-region Shfelat Yehuda.
'We're left with a wonderful territory in Shfelat Yehuda.'
(924)

Syntactic structure: NP ADJ DatNP (structure 28)

(7) *ha-uvdot yeduot li.*
the-facts known to.us.
'The facts are known to us.'

(5,446)

Syntactic structure: NP INTRANS.V DatNP (structure 32)

(8) *kama peamim* ***ze*** ***histovev*** ***li*** *b-a-rosh ad she-amarti OK.*
few times this turned around to.me in-the-head until that-I.said OK.
'I thought about it for a while till I said, OK.'
(4,679)

Syntactic structure: NP INTRANS.V DatNP A (structure 33)

(9) *karati et ha-dox ve-****hu nire*** ***li*** ***recini*** *beyoter.*
I.read ACC the-report and-it seems to.me serious very.
'I've read the report, and it seems very serious to me.'
(10,712)

Syntactic structure: NP TRANS.V DatNP (structure 41)

(10) *im* ***X ve-ozrav*** *lo* ***ozrim lanu*** *az anaxnu lo yexolim lacet*
if X and-his.assistants not help to.us then we not can exit
mi-sham.
of-there.
'If X and his staff won't help us we won't be able to get out from there.'
(1,184)

Syntactic structure: NP TRANS.V DatNP CL (structure 42)

(11) ***ani mavtiax lexa*** ***she****-anaxnu nisgor et ha-kaduregel.*
I promise to.you that-we will.close ACC the-football.
'I give you my word, we will shut down the football games.'
(113)

Syntactic structure: NP TRANS.V DatNP NP (structure 43)

(12) *eten laxem dugma.*
give to.you example.
'I'll give you an example.'
(141)

Syntactic structure: NP TRANS.V DatNP P (structure 44)

(13) ani mode lexa al ze.
I thank to.you on this.
'I thank you for this.'

(504)

Syntactic structure: NP TRANS.V DatNP V (structure 45)

(14) ***ten lo*** ***lehashlim*** *et ha-dvarim.*
give to.him to complete ACC the-things.
'Let him finish.'
(491)

Interpreting the associations presented in Figure 4.1, two conclusion can be drawn. First, we can see that there are three sets of Dative functions which occur with the same set of syntactic structures. That is, these sets of functions are related to the same construal of events encoded in these particular structures:

1. The Recipient, the Possessive Dative, and the Affectee occur with syntactic structure (43) more than expected.

2. The Addressee is the only function that occurs with syntactic structure (42) more than expected.

3. The Evaluative Dative and the Experiencer tend to occur with the same set of adjectival structures: (1, 3, 5, 28).

4. Contrary to other functions, the Experiencer cannot be easily paired with with a single syntactic structure or a single lexical category of predicate. In this respect, it fits its definition in the previous chapter as an umbrella covering many types of participant roles.

Second, as can be seen in Table 4.1, most of the Dative functions in the corpus occur with only one or two constructions with more-than-expected frequency, except for the Experiencer and the Evaluative Dative which span over a large set of constructions. That is, for the most part, a Dative interpretation is linked to a narrow range of syntactic environments. Moreover, this narrow range can be lexically/syntactically defined: The Recipient, the Affectee, and the Possesive Dative occur with a three-argument construction that includes a transitive predicate and a nominal complement (structure 43 in figure 4.1). The Discourse marker and the Human Endpoint tend to occur with the same construction; a two-argument construction with a Dative-marked direct object (structure 41 in figure 4.1). A variant of this type of constructions is relevant only for the Human Endpoint: a two-argument structure with an Infinitival complement of the predicate (structure 4). That is, it is a structure with two nominal arguments and a complex verb which consists of a Verb+Infinitive (resembling causative constructions

cross-linguistically, see von Waldenfels (2012)). The Addressee occurs with another variant of this structure: two nominal arguments and a finite clause as a verb-complement rather than an infinitival clause (structure 42). And finally, the Experiencer and the Evaluative Dative occur with Adjectival and Intransitive constructions, mostly with a Predicate-first syntactic order (1, 2, 3, 5, 22, 28). The Ethical Dative tends to pattern with the Experiencer, occurring in structure 32. That is, based on the interpretation of the Cohen-Friendly Association plot we can conclude that there are four form–function correlations in the data, but these correlations do not conform to traditionally defined Argument Structure Constructions; rather, they seem to conventionally pair a set of Dative functions with a set of syntactic structures:

(15)
a. The Recipient, the Affectee and the Possessive Dative share a pattern,
b. the Addressee has a pattern on its own,
c. the Human Endpoint and the Discourse Marker share a third pattern,
d. the Evaluative, the Experiencer, and to some extent the Ethical Dative share a fourth pattern.

In conclusion, using a Cohen-Friendly Association plot we have tested the hypothesis that it is syntactic structure, rather than semantically defined participant roles, which is the right solution for the Dative's distribution and behaviour in Hebrew. We have seen that some Dative interpretations are associated with a *pattern of uses*, rather than with a single syntactic structure which could be defined as a Construction. Thus, the answer should be sought for elsewhere – in a domain which concerns patterns among different syntactic structures, based on other parameters as well. This leads us to the next section which accounts for the data in a multivariate way. We will also introduce the notion of Discourse Profile Constructions in section 4.2.

4.2 Converging functions: Discourse profile constructions

After defining the Hebrew Dative's functions with regard to participant roles in chapter 3, showing that this approach is not sufficient for accounting for the corpus data, the next step was done in section 4.1: widening the point of view to include Argument Structure Constructions. However, while Argument Structure Constructions may account for some of the Dative usage patterns in the corpus,

such as the Addressee or the Recipient, it is not enough to account for the entire range of interpretations associated with the Hebrew Dative in the corpus. I therefore suggest in this section a multivariate discursive approach in order to account for patterns in the database corresponding to the whole range of Dative functions. This approach is multivariate in that it takes into account multiple parameters simultaneously, from multiple tiers of linguistic and extra-linguistic information, calculating each parameter's relations to the other parameters, across the entire database.

Thus, the next step in the investigation into the Hebrew Dative's usage patterns in the present corpus is carried out through exploratory multivariate statistics tools: Correspondence Analysis (CA, Baayen (2008); Glynn (2012)) in section 4.2.1, Multiple Correspondence Analysis (MCA, Rencher (2012)) in section 4.2.2, and Hierarchical Clustering on Principal Components (HCPC, Husson et al. (2011)) in chapter 5. Using these tools one can collect bottom-up information from the data and interpret patterns of associations between variables or between tokens of use (see section 2.3 for a methodological presentation).

4.2.1 Correspondence Analysis

In the previous section we have reached the conclusion that there are four patterns of use, two of them covering a range dative functions. This conclusion was based on the interpretation of the Cohen-Friendly association plot which considers deviations from expected frequencies of associations. Since there was no clear form–function correlation to be found for every case and every Dative function, this approach is now replaced with a multivariate approach, to be presented in the current and the following sections. However, the conclusion that there exist four usage patterns converging together several Dative functions is real enough, and must not be dismissed. We now turn to a different approach in order to provide further justification for the consideration of several functions under a single pattern, thus arguing for a similarity between these functions. First, we perform a Correspondence Analysis, to find out which syntactic structures correspond with each other, and which Dative functions correspond with each other. Finding a set of corresponding functions that correlates with a set of structures may reveal usage patterns of the type hypothesized in (15) above.

Let us define what it means to 'pattern together' when interpreting a Correspondence Analysis. In CA, two columns (or rows) that have similar profiles will be plotted closer to each other than two columns (or rows) that have different profiles. That is, in our case, two Dative functions that profile together in the sense that they participate in the same set of syntactic structures will be plotted closely

together. Thus, the CA bi-plot aids us in finding Dative functions that pattern together to the extent that they are used in the same set of syntactic structures.

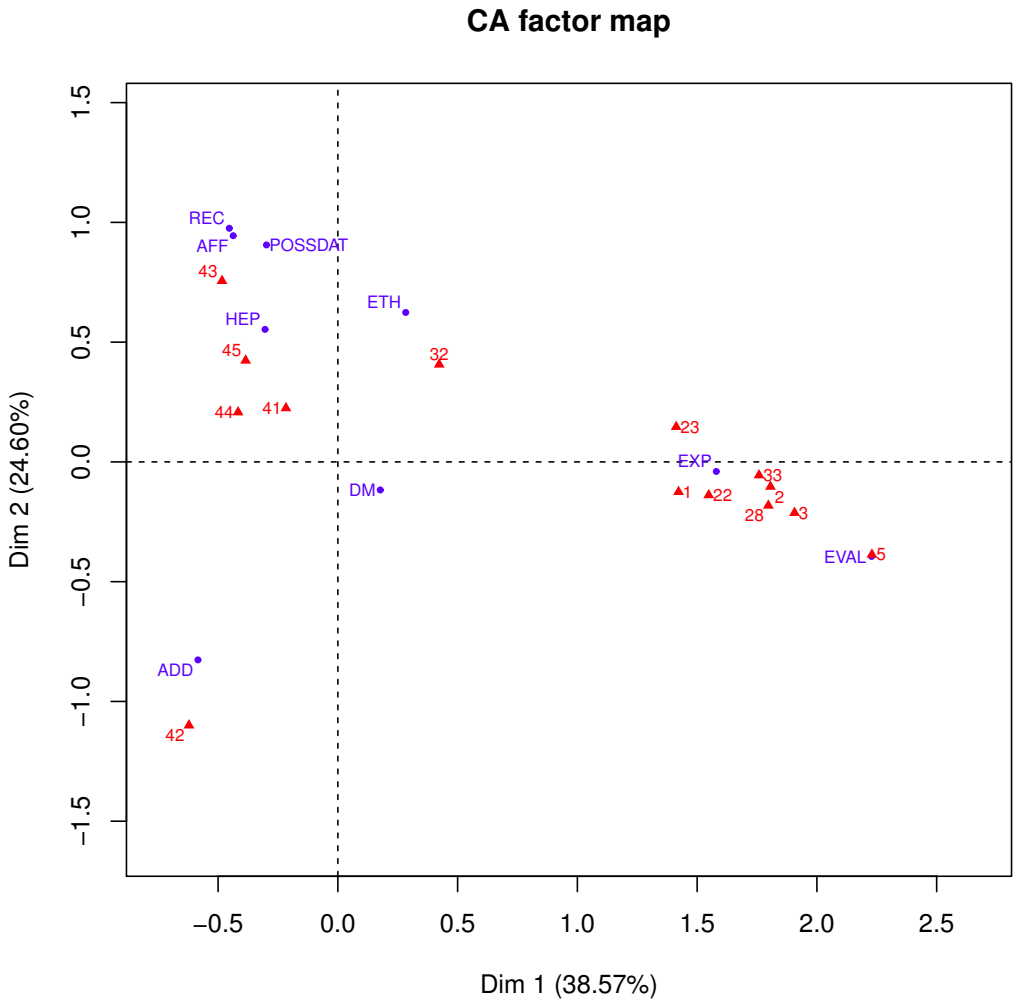
Figure 4.2 is a Correspondence Analysis of syntactic structures and Dative interpretations in the corpus. Indeed, we can see that the same patterns argued for above can be detected again. Table 4.2 is a cross-tabulation of syntactic structures and Dative functions in the corpus, presenting only the fourteen most frequent structures (all structures with a frequency larger than 1%, covering 95% of the data). This table is the input for the Correspondence Analysis presented in figure 4.2 (and for the association plot in Figure 4.1).

Table 4.2: Cross tabulations of syntactic structure and Dative function: 14 most frequent structures

		Dative functions								
	Argument structure	ADD	AFF	DM	ETH	EVAL	EXP	HEP	POSSDAT	REC
1	ADJ DatNP	1		49		39	125	12		
2	ADJ DatNP CL	2				31	335			
3	ADJ DatNP NP	2				27	70			
5	ADJ DatNP V	1				215	108			
22	INTRANS.V DatNP CL	15					166		1	
23	INTRANS.V DatNP NP		2	2	9		158	2	19	1
28	NP ADJ DatNP	5	1			79	126	14		2
32	NP INTRANS.V DatNP	49	7		19	1	205	153	6	101
33	NP INTRANS.V DatNP A						128			
41	NP TRANS.V DatNP	220	7	107	4		67	425	6	101
42	NP TRANS.V DatNP CL	2,246	2	31			12	35	2	
43	NP TRANS.V DatNP NP	539	135		15		9	425	173	1,433
44	NP TRANS.V DatNP P	62			1		1	147	1	
45	NP TRANS.V DatNP V	119					3	541		1

In figure 4.2 we see a two-dimensional mapping of the correspondences between Dative functions and between syntactic structures as presented in Table 4.2. Interpreting the bi-plot we first consider the correspondences between Dative functions (correspondence between columns), and then evaluate the correspondence between syntactic structures (row correspondence). Regarding Dative functions, dimension 1 (the X axis) corresponds to a distinction between event-related Dative functions on the left (Recipient, Possessive Dative, Affectee, Human Endpoint, Addressee), state-related Dative functions on the right (Experiencer, Evaluative Reference Point), and two mixed functions in between: Ethical Dative and Discourse Marker. The second dimension (the Y axis) corresponds to a

Figure 4.2: CA: Dative functions and Syntactic structure



distinction between affected participant on the top (Recipient, Possessive Dative, Affectee, Human Endpoint, and Ethical Dative), non-affected participant on the bottom (Addressee), and in between: Discourse Marker, Experiencer, and Evaluative Reference Point. Note that the Discourse Marker is adjacent to the $(0,0)$ point, emphasizing its neutrality as to affectedness and eventuality.

Considering syntactic structure we can see that the first dimension corresponds to the lexico-syntactic transitivity of the structure:[5] a distinction between three-

[5]in order to simplify the plot, argument structures are represented by numbers. See Table 4.2

argument constructions and transitive two-argument constructions on the left (structures 41–45 in Table 4.2), and intransitive constructions on the right (structures 1–33 in Table 4.2). The second dimension is harder to interpret with regard to the syntactic structures, since most of the points are placed in a dense cloud around the X axis. One distinction can be made, however, which emphasizes the difference between a nominal three-argument construction on the top-left and a clausal three-argument construction on the bottom-left.

These correspondences reveal three sets of Dative functions that pattern together. That is, in Figure 4.2 we can see three closely plotted sets of Dative functions: (1) the Recipient-Affectee-Possessive (and to some extent the Human Endpoint) on the top-left, (2) the Addressee on the bottom-left, (3) and the Evaluative-Experiencer on the right. The Ethical Dative and the Discourse Marker do not pattern with either of these sets. These sets of functions resemble the patterns found in the Cohen-Friendly Association Plot presented in (15) above, repeated in (16), with a slight change regarding the Human Endpoint:

(16)
a. The Recipient, the Affectee, the Possessive Dative, and to some extent the Human Endpoint share a single pattern,
b. the Addressee has its own pattern,
c. the Evaluative, the Experiencer, and to some extent the Ethical Dative manifest a third pattern.

That is, providing an argument from a different perspective, the Correspondence Analysis further supports accounting for the Dative functions as sets of related functions rather than isolated participant roles. This support, however, raises another question in turn. If indeed there exist such patterns, what else is shared between the tokens participating in these patterns, having already said above that considering only syntactic structure as a shared feature does not provide us with an adequate answer. The answer to this question can be answered only by considering multiple parameters at once. This can be done by an extension of the Correspondence Analysis method, namely, Multiple Correspondence Analysis. This is the topic of the next section.

4.2.2 Multiple Correspondence Analysis

Syntactic structure, we have seen in previous sections, is not an adequate organizing principle for the distribution and usage patterns of Hebrew Dative functions. However, I propose that an examination of multiple parameters distinguish be-

for a reader of constructions' numbers.

tween the various functions associated with the Dative in Hebrew – discursive differences, rather than merely semantic ones. That is, in order to identify such differences we ought to examine each Dative token of use in a wider context; we need to analyse the relations between a Dative function and a discourse pattern. This discourse pattern should include as many parameters as possible, covering multiple tiers of linguistic and extra-linguistic information.

Such an inquiry requires a fresh point of view on corpus data, taking into account multiple parameters simultaneously. Exploratory multivariate statistics provides us with tools for such a consideration of the data. Multiple Correspondence Analysis (MCA) is the main technique used in the present study. As was presented in section 2.3, MCA points to tendencies within a large, multivariate database, and sheds light on shared and unique parameters within every set of similar data points – similar tokens of use or similar categories. Using MCA, for example, one can tell why a particular token of use is judged as similar to another token, or less similar – or different – from yet another token of use. Thus, we can detect sets of similar tokens – a cluster of exemplars – and ask whether they conform to a constructional definition; that is, whether they constitute an emergent construction that conventionally links some sort of form to some sort of function.

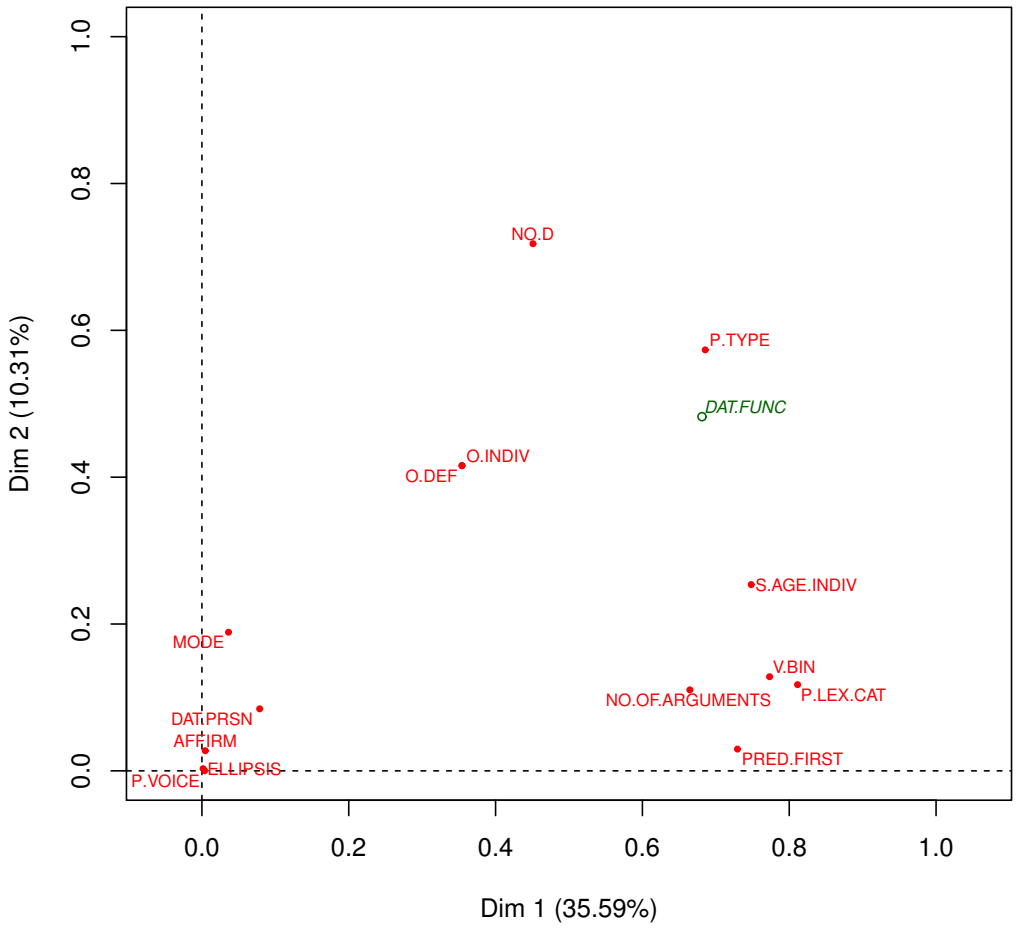
Since we consider multiple parameters, from multiple sources of information, the conventional form-function correlation we define using MCA can be wider than a syntactic-semantic pairing, and includes lexical, morphological, pragmatic and discursive parameters on top of the usual candidates of syntax and semantics. This type of enriched constructional context is referred to here as a *Discourse Profile Construction*; a wide context unit of formal and functional parameters that conventionally co-occur in the language, thus constituting a cognitive category.

Regarding discourse patterns as constructions is a unique perspective on constructions in general and on Argument Structure Constructions in particular. Thus, Constructions in the framework advocated here, are an amalgam of lexical, morphological, syntactic, semantic, pragmatic, and discursive information – at least. Examining the present corpus from such a point of view we can see that each Discourse Profile Construction corresponds, to a certain extent, to a traditionally defined Dative function, or participant role. That is, such an analysis suggests new Dative functions, based on the Discourse Profile Construction they participate in.

I calculated a Multiple Correspondence Analysis (MCA) in order to find out hidden correspondences and patterns in the database as candidates for an organizing principle of the Hebrew Dative usage. The procedure of performing an MCA was spelled out in section 2.3. A numerical summary of the results is provided in appendix B.

Figure 4.3 presents the dimensions of the map, its axes, relative to fifteen

Figure 4.3: Multiple Correspondence Analysis: contribution of parameters to the first two dimensions



parameters with a combination of which each token of use can be characterized. Note that the dimensions of the MCA are constructed according to a set of active variables defined by the researcher. Only those variables I have hypothesised as candidates for explaining the distribution of Dative interpretations are included as active variables. That is, the Dative function parameter itself, although coded in the corpus, was left out of the construction of the MCA map. Had this variable been included in the calculation the argument would of course be circular.

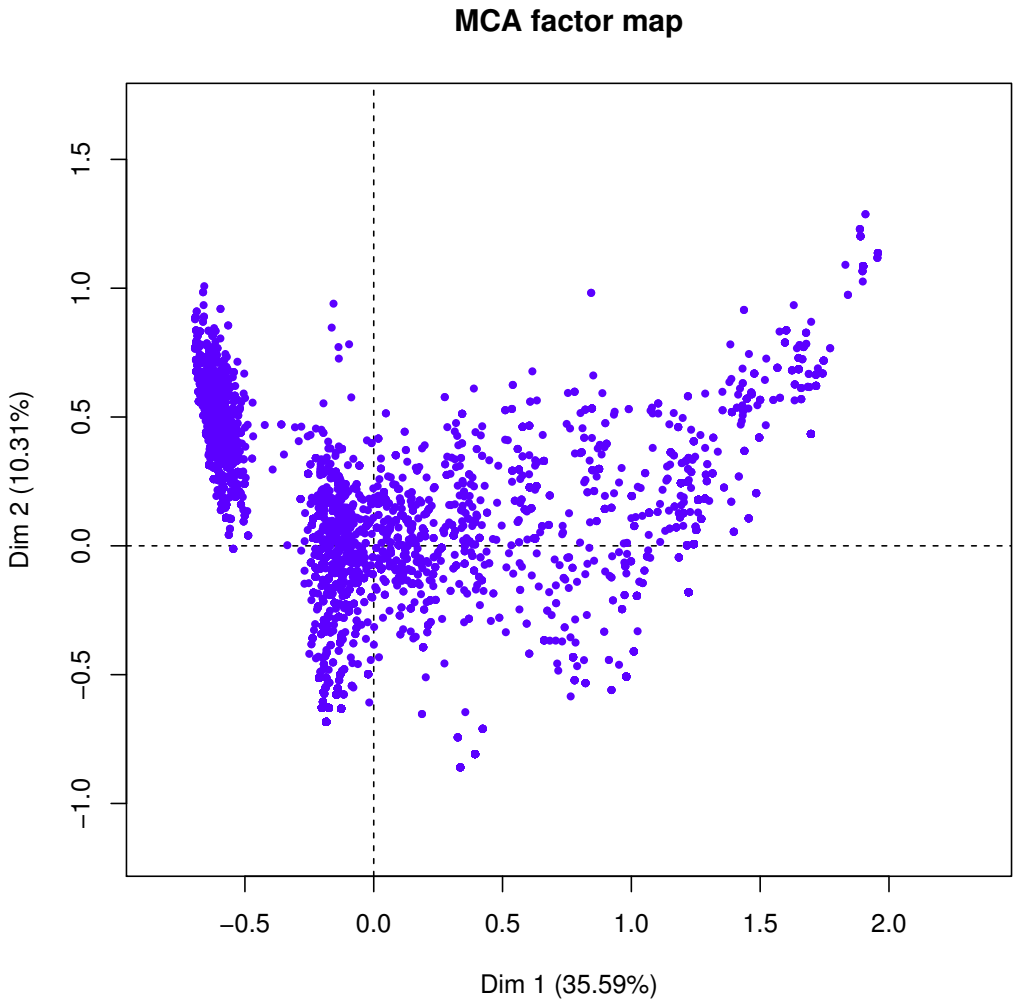
Figure 4.3 provides a description of the strength of the link between a cat-

egorical variable and a dimension, such that the higher the variable is located with regard to a dimension, the stronger the link. We can see that the variable *P.Type* is strongly linked to both the vertical and the horizontal dimension, meaning that two tokens placed far from each other — vertically or horizontally — cannot be characterized by the same semantic type of predicate. Other variables are linked to either one of the dimensions: the horizontal dimension is linked to features of the Subject argument and the predicate: the agency of the Subject referent (S.AGE.INDIV), the verbal morphological paradigm (*Binyan*, V.BIN), the lexical category of the predicate (V.LEX.CAT), the number of arguments in the clause (NO.OF.ARGUMENTS), and Subject-Predicate order (PRED.FIRST). That is, sets of tokens situated apart from each other on the horizontal dimension do not share the aforementioned parameters. The vertical dimension is linked to features of the Direct-Object argument: the type of non-Dative argument in the clause (NO.D), the level of individuation of the Direct-Object referent (O.INDIV), and the definiteness of the Direct-Object argument. The variable concerning realis vs. irrealis clauses (MODE) is very weakly linked to the second, vertical dimension. That is, sets of tokens situated apart from each other on the vertical dimension cannot be characterized by the same combination of these variables. The variables *person* of the Dative referent (DAT.PRSN), *affirmation* of the clause (AFFIRM), *ellipsis* (ELIPSIS), and *voice* (whether the predicate is in the passive or middle voice (V.PAS.MID)), are not linked to either of the dimensions, meaning that contrary to the parameter *P.Type*, two distant sets of tokens (vertically or horizontally) cannot be distinguished by one of these parameters.

A second perspective on the Multiple Correspondence Analysis map is the correspondence between individuals, or in our case, between tokens of use. In order to understand these relationships we need to study the similarities between tokens, calculated over each and every variable. Through such a perspective we can ask, for instance, which tokens are the most similar, what makes them similar, and are there groups of similar tokens. According to the Multiple Correspondence Analysis, the degree of similarity between two tokens is determined according to the number of parameters they share (but see section 2.3 for some exceptions). That is, two tokens are considered highly similar if they share the same category for each of the variables. The distance (dissimilarity) between tokens is thus calculated category by category, together with a consideration of the frequency (high or low) of the category in the entire corpus (Husson et al., 2011). Consider, in this respect, Figure 4.4, in which the cloud of individuals in the sample is presented. The cloud of individuals is a representation of all the tokens in the data, arranged as to degree of similarity.

On a first glance, we can see some sort of order in the distribution of tokens

Figure 4.4: Multiple Correspondence Analysis: cloud of individuals



on the map, with dense sets of tokens to the left side of the Y axis, and scattered groups of tokens populating the right side of the Y axis: (1) A dense group of tokens is placed in the top-left quadrant, (2) another dense group is situated around the origin, (3) a third group, a sparse one, is situated in a portion of the bottom-left quadrant, and (4) a large group of tokens is spaciously scattered on the right side of the map. So far this indication cannot tell us much more than the general claim that there is a skewed distribution, with two groups of tokens sharing many features ((1) and (2)), one small group (3), and a large group of tokens which

share some parameters among themselves, but share very few features (if any) with other groups of tokens (4).

In order to better interpret the MCA map we need to consider both the contribution of each parameter to the construction of the map, and the distribution of the parameters' levels across the cloud of tokens. These steps will lead us to a better understanding of the clusters of tokens visible in Figure 4.4. It is important to note at this point that the current superficial glance on Figure 4.4 will receive a statistical support by using Hierarchical Clustering on Principal Components in chapter 5, where we will see that these groups of tokens are indeed clusters with high degree of within-group similarity and across-group dissimilarity. First, however, we must better understand the MCA map. This is done by interpreting Figures 4.3–4.13.

4.2.2.1 Type of predicate

The first variable that is potentially correlated with the interpretation of the Dative-marked participant is the semantic type of the predicate. FIgure 4.6 shows the correspondences between the eight semantic types, and in Figure 4.5 we can see the distribution of the tokens in the data according to their respective type of predicate. Note that the coordinates of each type of predicate (its location on the map) is a function of its distance from every other category of the same variable. Simply put, this distance is calculated by counting the individuals that carry either of the categories (but see (Husson et al., 2011) for a complete explanation of this calculation).

Starting from the right side of the map, the type marked as IRR refers to tokens with nominal predicates. We can see that they are located on the far right side of the map, distant from the other tokens. That is, they share very little features with other Dative clauses.

A step to the left and downwards, we find the adjectival types of predicates: Property and Value. The semantic types Property and Value are adjacently located at the right half of the map, meaning that the tokens involving these two types share other functions. These two semantic types refer to tokens such as the following examples:

(17) Property:

a. *koev li ha-lev al ma she-kore b-a-maxon.*
hurts to-me the-heart on what that-happens in-the-institute.
'My heart aches for what's happening in the institute.'
(11)

Figure 4.5: Multiple Correspondence Analysis: tokens coloured according to type of predicate

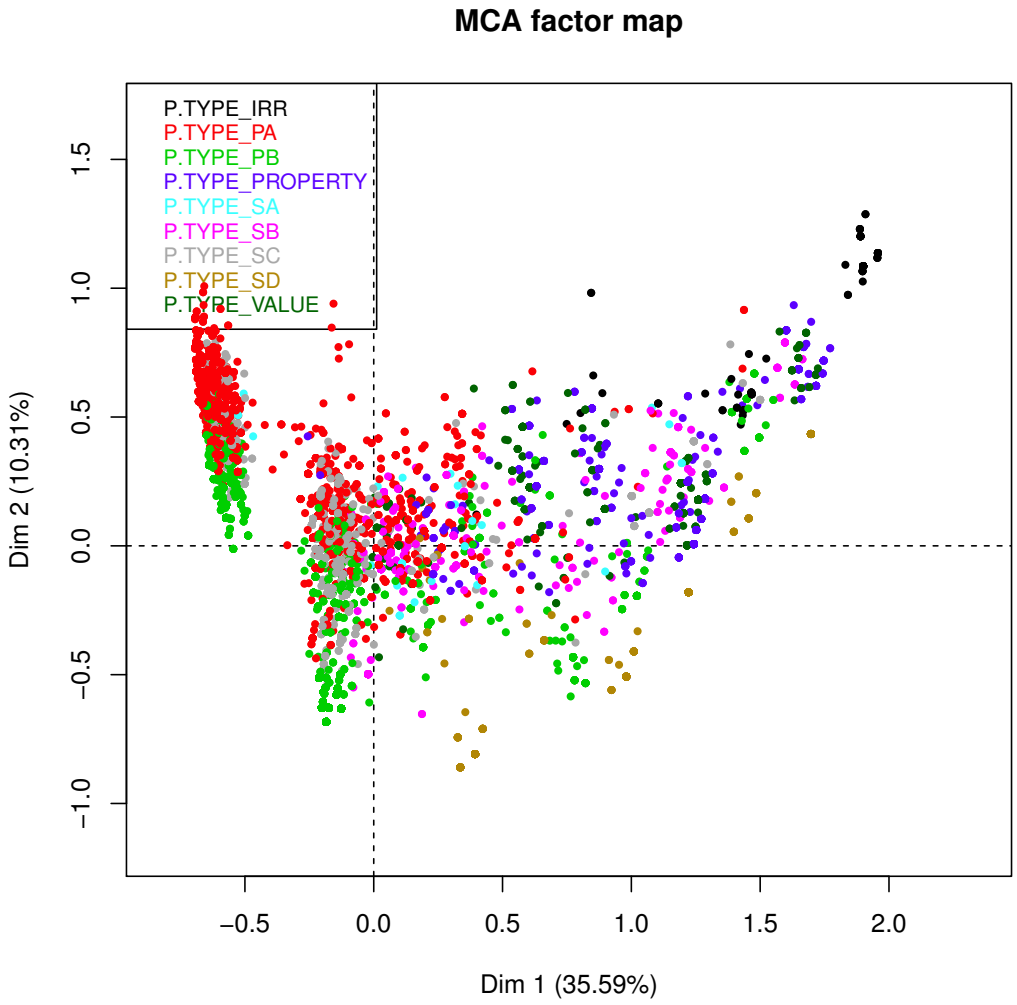
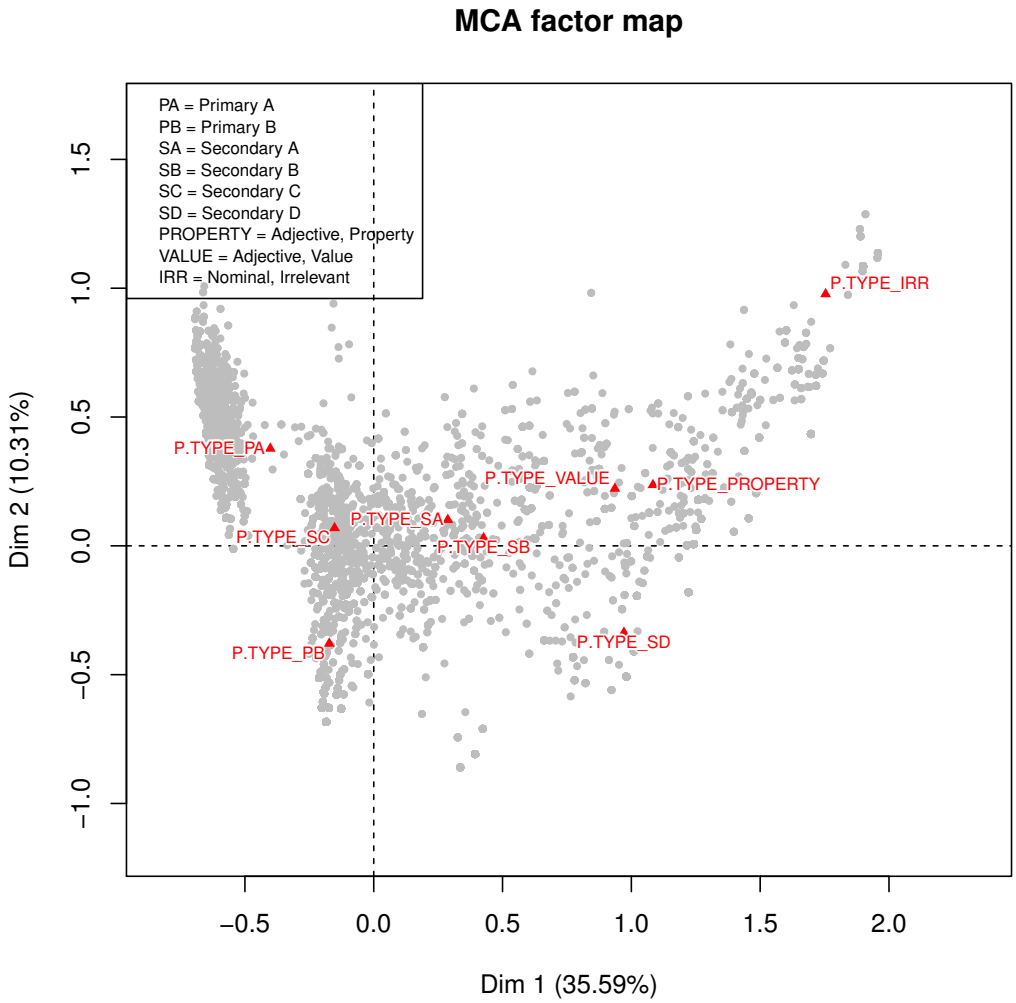


Figure 4.6: Multiple Correspondence Analysis: correspondences between semantic types of predicate



b. *lo naim li lehalot etxem be-teurim lo neimim.*
not pleasant to-me to-weary you in-descriptions not pleasant.
'I feel bad with wearying you with unpleasant descriptions.'
(975)

(18) Value:

a. *asur lanu ke-xavrey-kneset lihyot shutafim*
forbidden to-us as-Knesset-members to-be partners
l-a-bizayon.
to-the-humiliation.
'We, as parliament members, shouldn't be part of this humiliation.'
(740)

b. *xashuv li lomar zot be-ofen bote.*
important to-me to-say it in-matter harsh.
'It's important for me to say it harshly.'
(1036)

The semantic types Value and Property depict no effect on the Dative-marked participant, and are lexicalized through Adjectives. They construe a situation relative to the Dative referent, personalizing it. For instance, a non-Dative version of (17a) would result with an impersonal utterance:

(19) *koev ha-lev al ma she-kore b-a-maxon.*
hurts the-heart on what that-happens in-the-institute.
'The heart aches for what's happening in the institute.'
(∼(17a))

"The heart aches" is the reported situation in both (17a) and 19, the difference between the two being the possession relation in (17a) that relates the heart-ache to the Dative-marked participant. It is not the case (17a) construes the result of whatever happened in the institute. Rather, the situation 'the heart aches' is reported. Had the effect of what happens in the institute been the desired construal, the sentence should have been:

(20) *ma she-kore b-a-maxon maxiv li / shover li et*
what that-happens in-the-institute hurts to.me / brakes to.me ACC
ha-lev.
the-heart.
'What's happening in the institution hurts me / brakes my heart.'
(∼(17a))

Below these adjectival types is the Secondary D type. The predicates belong-

ing to the Secondary D type, as presented in section 2.1.2, are intransitive verbs with complement clause in S relation. In our case, these verbs take another role additional to those of the complement clause: the Arbiter, marked by the Dative. That is, secondary D type predicates do not profile an affected entity. Rather, they convey a subjective stance or belief regarding the situation described in the clause in A position. For example:

(21) *nidme li she-sarfu sham beyt sefer.*
seems to.me that-burnt there school.
'I think a school was burnt over there.'
(7,895)

The secondary A and secondary B types are located adjacently, meaning that the tokens characterized by these predicates share other features. Both secondary A and B modify the meaning of another verb. However, while secondary A predicates share participants with the situation they modify, secondary B predicates have an independent participant (either the Dative-marked participant in two-argument structures, or the A participant in three-argument structures) whose attitude towards a situation is described. The Dative-marked participant role in such clauses varies, and cannot be uniformly characterized. For example, the following sentences all share a Secondary A/B predicate, yet their Dative-marked participant does not always bear the same function:

(22) Secondary A:

a. *kedey lehaxmir et ha-anisha* ***lo maspik lanu*** ***takanot***.
in order to worsen ACC the-punishment **not enough to.us** **regulations**.
'If we want harder punishment, regulation are not enough for us.'
(12,451) Experiencer

b. *anaxnu shoafim lehafsik lo et ha-tipul.*
we aspire to stop to.him ACC the-treatment.
'We'd like to stop his treatment.'
(14,669) Possessive Dative

(23) Secondary B:

a. *ma ixpat li?*
what care to.me?
'What do I care?'
(8,380) Experiencer

b. *nimas li kvar mi-kol ha-diburim.*
fed up to.me already from-all the-talking.
'I'm fed up with all this talking.'
(6,643) Experiencer

c. *hayiti mamlic laxem livdok et ha-hodaot*
I.would.have recommend to.you to check ACC the-messages
she-atem mociim l-a-tikshoret.
that-you post to-the-media.
'I would have recommend you to check your announcements to the media.'
(2,661) Addressee

Considering these examples, we can see that there is nothing in common between the Dative-marked participant of clauses with secondary A or B types of predicates: it may be affected as in (22b), an experiencer of a feeling (23a), (23b), or the audience of a message (23c). Indeed, their location at the middle of the map, near the origin point, suggests that they are not related to any particular or unique construal. That is, the middle of the map is populated with tokens that may share features with both the tokens on the right and tokens on the left. As we go on interpreting the Multiple Correspondence Analysis map, we will better understand the significance of of this point. First, however, we should return to describing semantic types, moving to the right half of the MCA map.

The secondary C type of predicates is located in the top-left quadrant, very close to the origin point. Like all secondary verbs, secondary C verbs modify another situation. However, in secondary C verbs the participant introduced by the verb (on top of the participants of the complement verb) plays an active role in bringing about the situation described by the complement clause or Direct Object. For example:

(24) a. *carix lidog lahem le-diyur savir.*
need to worry for.them for-accommodation reasonable.
'We have to supply them with reasonable accommodation.'
(6,803)

b. *kvucot ha-kaduregel ifsheru lanu knisa xofshit l-a-migrashim.*
football teams enable to.us entrance free to-the-fields.
'We've been allowed by the teams to go into their games free of charge.'
(1,220)

Also in the top-left quadrant are tokens with predicates of the primary A type,

concerning actions with concrete participants:

(25) a. *ha-tokbak ha-yexidi ha-ze horeg oto, mexasel oto,* ***gomer***
the-talk-back the-single the-this kills him, destroys him, **finish**
lo et ha-xayim, mefarek lo et ha-mishpaxa.
to.him ACC **the-life**, **disassemble to.him** ACC **the-family**.
'This single talk-back kills him, destroys him, finishes his life, tears his family apart.'
(13,981)

b. *ne'esham she-eyn lo orex din* ***beyt ha-mishpat***
defendant that-there.is.no to.him lawyer **the court**
ma'amid lo orex din.
places to.him orex din.
'A defendant that has no lawyer, the court provides him with one.'
(9,341)

Example (25b) provides an interesting addition to the debate with regard to the differences between the Affectee and the Possessive Dative (see, for example, Landau, 2010). On its own, the Dative-marked participant in the second part of (25b) is a beneficiary of the court's actions, an Affectee. The lawyer provided by the court is not in his possession, thus the Dative-marked participant should not be interpreted as a Possessive Dative. However, consider the first part of (25b): here the possession relation between the defendant and a lawyer (or the lack of it) is explicit. The second part of the of the sentence describes a solution for the problem described in the first part. Since the problem concerns lack of possession, the solution must include granting possession. That is, contrary to the conclusion above, the Dative-marked participant must be interpreted as a Possessive Dative. This contradiction strengthens the claims by Linzen (2009) and Boneh and Bar-Asher Siegal (2014) that the Possessive Dative and the Affectee roles cannot be separated. This claim will receive further strengthening as we analyse the clusters in the data using Hierarchical Clustering on Principal Components.

The last conceptual type of predicates, the primary B type, is in the bottom-left quadrant. This type also concerns concrete participants, but can be complemented by a clause as well:

(26) a. *nimsar li she-gam ha-rofim lifamim xolim.*
delivered to.me that-also the-doctors sometimes sick.
'I've been told doctors can get sick as well.'
(40)

b. *histaber lanu she-eyn shinuy.*
turned out to.us that-there.is.no change.
'As it turns out, nothing has changed.' (10,845)

c. ***ani modi'a lexa she**-yesh lo od.*
I inform to.you that-there.is to.him more.
'I'm telling you, he has more.'
(1,324)

Considering the correspondences between predicate types, we can see that they are distributed on the Multiple Correspondence Analysis map according to the level of effect inflicted upon their Dative-marked participant. On the left there is the Primary A type, in which the Dative-marked participant is prototypically affected as a result of a concrete event. This is the highest level of effect that can be marked by the Dative in Hebrew. For example, consider the following examples:

(27) a. *holxim lishtot li et ha-maym me-ha-shorashim shel*
walk to.drink to.me ACC the-water from-the-roots of
ha-cmaxim.
the-plants.
'They're gonna drink my plant's water (= they are going to draw water from a nearby underground location, whereas my plants will loose their water supply).'
(882)

b. *axarey ze baim ve-shotim lahem et ha-sibsud*
after that (they).come and-drink to.them ACC the-subsidization
ha-ze im toxniot pratiot.
the-this with plans private.
'And then somebody will waste their subsidization with private plans.
(2562)

These examples show us that clauses with Corporeal predicates (although metaphorically used in (27b)) construe a strong effect on the Dative-marked participant as resulting from the Agent's action on the Direct Object referent. A similar effect can be detected in clauses with other Primary A type verbs:

(28) a. *lefi daati yikxu lexa et ha-rishayon.*
according my opinion they.will.take to.you ACC the-license.
'I believe they will confiscate your license.'
(332)

b. *anshey ha-mishtara gam heviu lanu doxot.*
people the-police also brought to.us report.
'The police gave us reports as well.'
(1199)

(29) a. *hecatem lehaxzir la 280 shekel.*
you.suggested to.return to.her 280 Shekel.
'You've suggested you'll give her 280 ILS back.'
(101)

b. *nisinu lehagish laxem et ha-mismax ha-ze.*
we.tried to.submit to.you ACC the-document the-this.
'We've been trying to submit this paper to you.
(527)

These two types share many parameters, in the sense that they construe a very similar state of affairs. In the following example we can see another Primary A verb. Using verbs such as 'create' (30a) or 'issue' (30b) below, the speaker construes a situation in which the creation of the Direct Object referent affects the Dative-marked participant in some way.

(30) a. *ze yicor lanu macav meod mesubax.*
this will.create to.us situation very complicated.
'It will cause a very complicated situation for us.
(563)

b. *nashim she-eyn lahen darkon crixot she-yanpiku*
women that-there.is.no to.them passport need that-will.issue
lahen teudat maavar.
to.them laissez-passer.
'Women with no passport need to be issued with a laissez-passer.'
(762)

One level of effect lower, and lower and to the right on the Multiple Correspondence Analysis map, is the Secondary C type. In Secondary C verbs the Agent actively brings about a situation, for example, by enabling or allowing:

(31) a. *ani lo doeg lexa.*
I not worry to.you.
'I don't worry about you.'
(6,421)

b. *ata marshe li ligmor mishpat?*
you allow to.me to.finish sentence?
'Will you let me finish my sentence?'

(11,559)

On the same level of affectedness we can find the Primary B type. Consider the following example:

(32) a. *ani mecia lexa ke-ohed she-tisgor oto le-reva shaa.*
I suggest to.you as-fan that-you.will.close it for-quarter hour.
'As a fan, I suggest that you should close it for fifteen minutes'.
(30)

b. *ani mavtiax lexa she-tipul b-a-vaada ihye mahir.*
I promise to.you that-treatment in-the-committee will.be quick.
'I promise you that the committee's treatment will be quick.'
(229)

The Agent in Speaking events (for instance, (32a)) is performing a two participant action of uttering something out-loud. In a way, it resembles concrete events in that the Agent is creating a suggestion. This utterance, however, has an effect on its audience only to the extent that they hear it and may act accordingly as a result. Contrary to concrete two-participant events, in which the creation of a physical object might affect a third participant by being transferred to its possession or benefiting him in some way, the effect transferred from the Agent to the Dative-marked participant in Speaking events is different. It is an abstract, psychological effect. With regard to an affectedness scale, then, the Speaking semantic type is located just above the Value and the Property, indicating low affectedness of the Dative-marked participant.

As we go down the map, the degree of effect is decreased as well. Consider the following examples:

(33) a. *todiu li ma kara.*
you.will.inform to.me what happened.
'Let me know what happened.'
(1,076)

b. *ani ar'e lexa mismax history.*
I will.show to.you document historical.
'I will show you a historical document.'
(2,450)

The predicates in (33) resemble the Speaking semantic type presented in (32), in that they all concern transferring information and they all bring about a psychological effect on the Dative-marked participant. They all involve an abstract effect on the Dative-marked participant, but they construe it in a different manner.

At the current stage of the analysis we cannot explicitly define the differences in construal between the two sets of examples. In order to do that we need more information concerning other parameters such as the type of the Direct-Object, introduces later on.

Going right on the map, we find two similar types: Secondary A and Secondary B. In these types, there is no Agent that brings about an effect on the Dative-marked participant:

(34) a. *lo ixpat lo.*
not care to.him.
'He doesn't care.'
(242) Secondary B

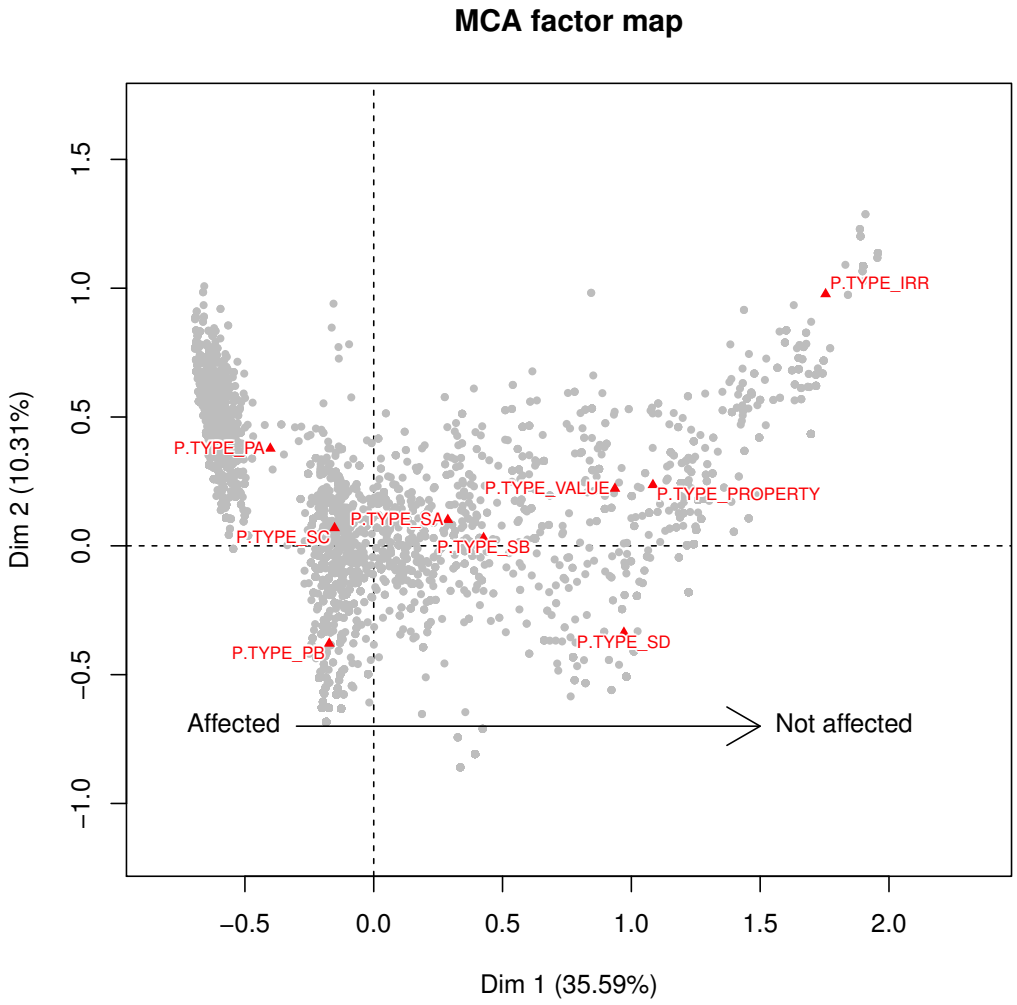
b. *lo maspik laxem ha-mehumot she-rainu be-london?*
not enough for.you the-riots that-we.saw in-London?
'Weren't the riots we've seen in London enough for you?'
(12,204) Secondary A

Further to the right we find the Adjectival types on the top-right quadrant of the map, and the Secondary D type on the bottom-right quadrant. Here, as in Secondary A and B, there is no effect that stems from an Agent's action. Rather, there is a subjective stance taking (Secondary D), and evaluation or experience of a situation. In all of these types, the Dative is not affected. Rather, it is a cognizer, a sensor, and arbiter, exemplified in (35):

(35) *keday lexa la'uf mi-po kama she-yoter maher.*
better for.you to fly from-here as that-more fast.
'You better run away as fast as you can.'
(79)

Summing up the discussion about the semantic type of predicates, we can see that the different types are distributed along the map according to the level of affectedness of the Dative-marked participant associated with them. That is, the first organizing principle for the Hebrew Dative category, according to which two Dative clauses can be compared and differentiated, is the semantic type of predicate. This is the Affectedness Scale, presented in Figure 4.7. As we go on to account for the other variables in the data, we will see that the distribution of categories according to an Affectedness Scale is a recurrent finding, thus strengthening its role as a cognitively relevant organizing principle for the category.

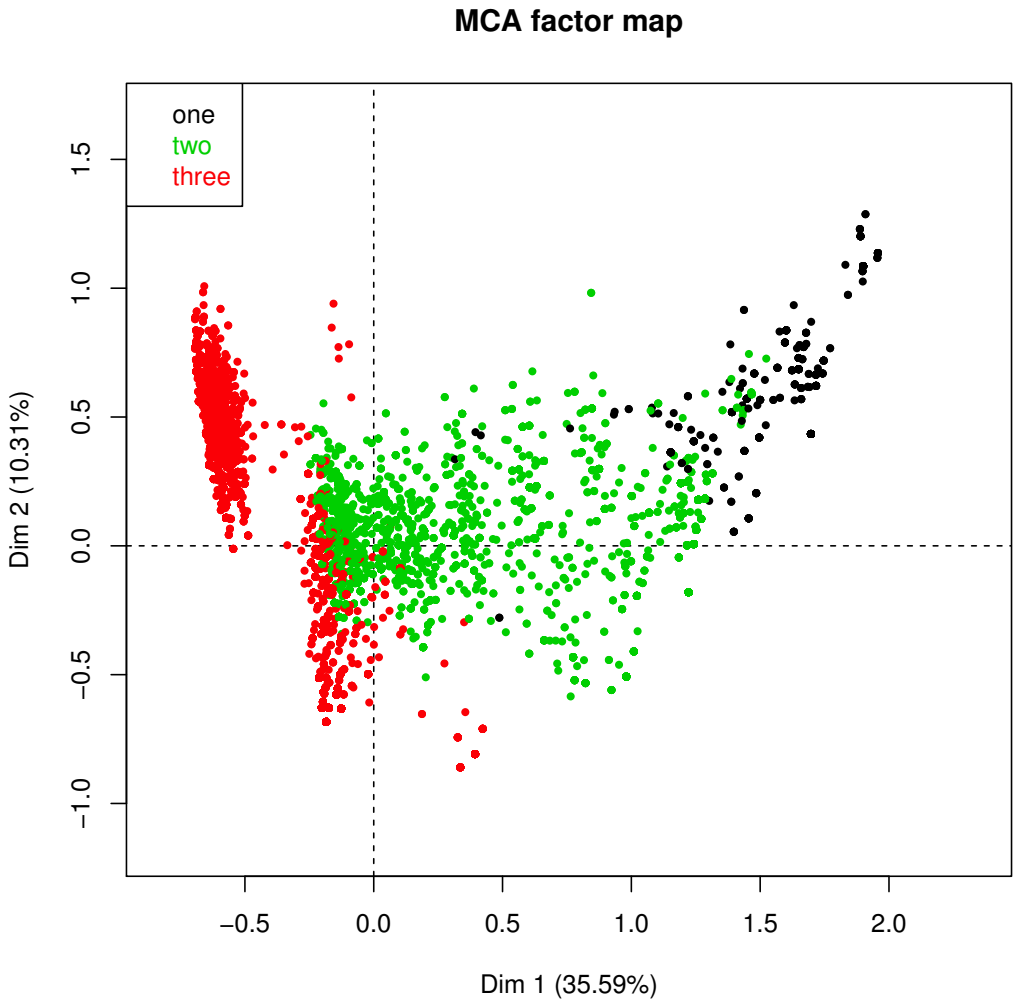
Figure 4.7: Affectedness scale: semantic type of predicates



4.2.2.2 Number of arguments in the clause

Each token in the corpus was coded for several syntactic features. The first syntactic variable to be considered here is the number of arguments in the clause, presented in Figure 4.8. Figure 4.8 presents a clear distribution of the tokens on the map according to the number of arguments in the clause of each token. This distribution is highly homogeneous: tokens with three-arguments are grouped together, and the same is true for tokens with one- and two-arguments. That is, the number of arguments in the clause is a good candidate for being an organizing

Figure 4.8: Multiple Correspondence Analysis: number of arguments in the clause



principle for the Hebrew Dative category. In the context of identifying an affectedness scale, the number of arguments is correlated with a particular type of effect. An effect on the Dative-marked participant that is mediated by another participant in the event is different than a direct effect on the Dative-marked participant, and these two are different from a null-effect. A mediated effect is construed by a three-argument construction (the Actor, the Affectee, and the mediating participant), a direct effect is construed with a two-argument construction (the Actor and the Affectee), and the null-effect is construed with a stative, one-argument con-

struction. Thus, we can consider the number of arguments to be an indicator for the Dative's interpretation with regard to the effect inflicted on the Dative-marked participant.

4.2.2.3 Type of the non-Dative argument

The second syntactic variable refers to the type of non-Dative argument; in three-argument structures, it is the element in Direct Object relation, while in two-argument structures, it is the second, non-Dative argument. In the following figures we can see an MCA map with tokens coloured according to their non-Dative argument type (Figure 4.9), and a map with correspondences between the types of non-Dative arguments found in the corpus (Figure 4.10). Recall that in Figure 4.3 on page 67 we have seen that the non-Dative type of argument parameter is linked to the second (vertical) dimension. Here we can see a manifestation of this link. Notice that the top half of the map is populated mainly with NP (Noun Phrase) and IRR (for irrelevant: one-argument structures), while the bottom half is occupied with CL (for a clausal complement), P (for Prepositions), and A (for Adjectives). The V type (for infinitival verb complement, or an infinitival clause), is located at the center of the map. That is, the type of verb-complement plays a major role in structuring the Dative category, clearly distinguishing between nominal complements and everything else.

However, even though the map is ordered in some way according to the non-Dative participant, this parameter cannot be claimed to be an organizing principle for the Dative category. That is, as we have seen so far, there is not much in common between tokens from the right half and the left half of the map. Thus this parameter—on its own—cannot function as a source for comparison between tokens of use. If so, let us add another parameter to the syntactic consideration of the clause, discussing not only the number of arguments, and type of the non-Dative argument, but also the linear relations between constituents. This is done by considering the distribution of the parameter *Word Order* on the MCA map, in the next subsection.

Figure 4.9: Multiple Correspondence Analysis: non-Dative argument

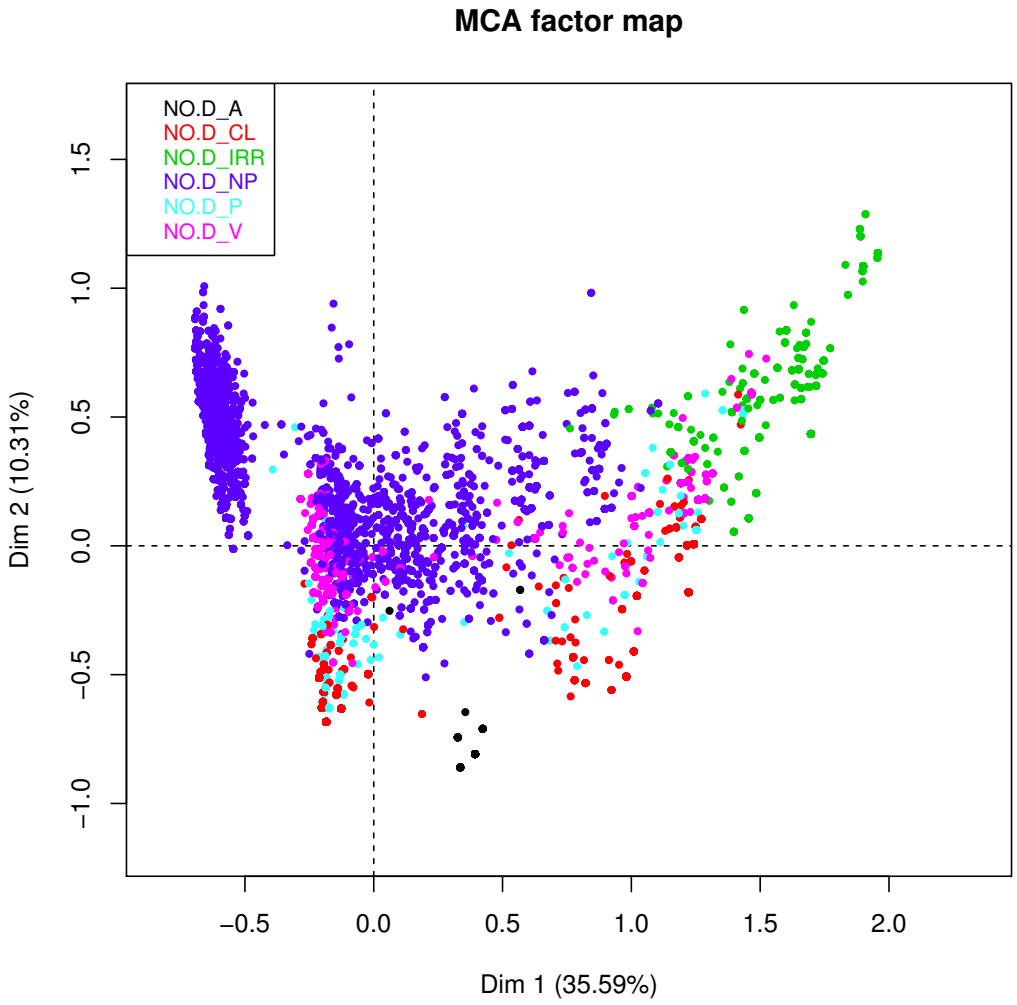
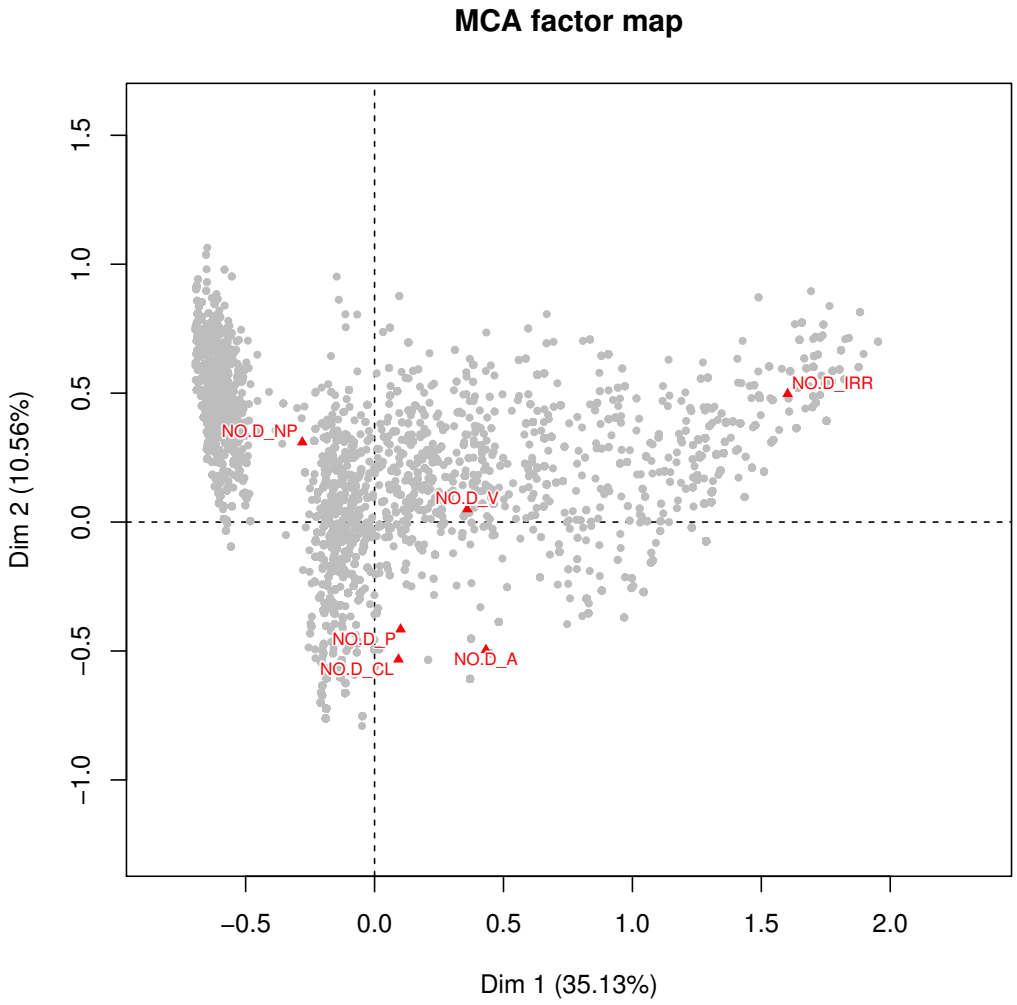


Figure 4.10: Multiple Correspondence Analysis: non-Dative argument II



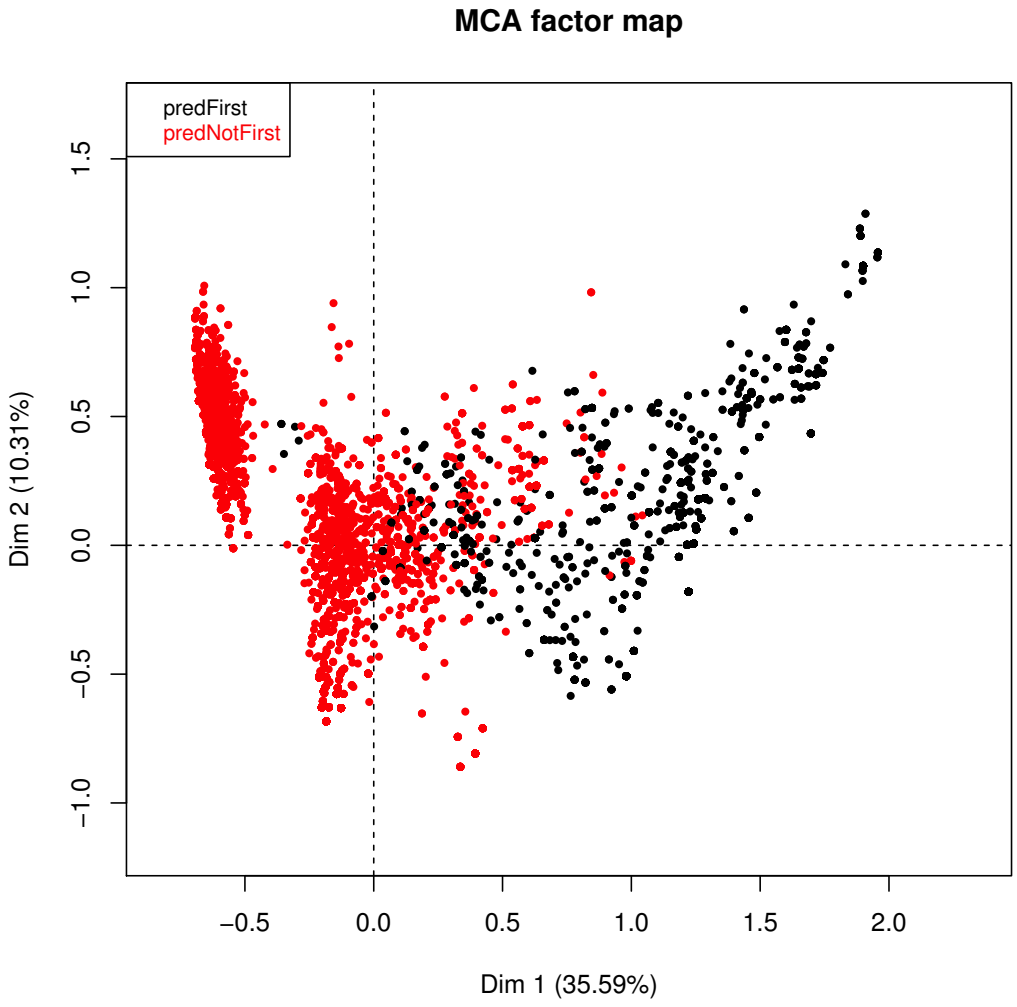
4.2.2.4 Word order

The Word Order parameter, or Subject-Predicate order,[6] as was presented in Figure 4.3 on page 67, has a strong link to the horizontal dimension of the MCA map.

[6] The Word Order variable was coded for the location of the predicate within the clause: [Predicate-First:yes/no]. However, this parameter is analysed here as encoding word order in general, rather than merely the location of the predicate. This is so since word order variants which included the entire argument structure have proved irrelevant for the analysis. These variants are related either to information structure parameters or to different types of emphasis strategies, while the Predicate-First parameter has proved to be related to the Dative interpretation. Thus, while a different course of research would have found a great interest in such word order variants, the

And indeed, Figure 4.11 shows that the Word Order parameter divides the cloud of individuals in half. As we go from the origin point left on the X axis, the word

Figure 4.11: Multiple Correspondence Analysis: word order



order is such that the predicate does not occur first in the clause, while from the origin point to the right on the X axis, the clauses begin with a predicate. This division of the map to right vs. left has been noted above. That is, accounting for the distribution of the Word Order parameter in the corpus in the context of previ-

current perspective may gain only from considering the linear relations between the predicate and the rest of the constituents in the clause.

ously discussed parameters we conclude that the clause's word order corresponds to a set of unrelated features such as the nature of the effect on the Dative-marked participant (see section 4.2.2.1 above) or the number of arguments in the clause (section 4.2.2.2), such that combinations of these variables form clusters of similar tokens of use. In other words, the clause's word order is another promising source for comparison between tokens for finding functional similarity (and dissimilarity).

We now expand our search for an organizing principle out of the semantic-syntactic domain, considering two related parameters: the lexical category of the predicate, and its morpho-syntactic paradigm.

4.2.2.5 Lexical category of the predicate

The predicate of a Dative clause belongs to one of nine lexical categories, according to the present corpus: a Transitive Verb (coded as TRANS.V), an Intransitive Verb (INTRANS.V), an Adjective (ADJ), and Adverb (ADV), an Noun (N), a Preposition (P), an Interrogative, a Discourse Marker (D.M), and a Complex Verb (COMP.V). (36)–(44) exemplify Dative clauses with each of these lexical categories, while their frequencies in the corpus are provided in Table 4.3.

Table 4.3: Frequencies of lexical categories of the Dative predicate

Type	**Tokens**	**Relative frequency**
Transitive Verb (TRANS.V, 36)	6,939	71.58%
Adjective (ADJ, 37)	1,293	13.34%
Intransitive Verb (INTRANS.V, 38)	1,242	12.81%
Noun (N, 40)	117	1.21%
Adverb (ADV, 39)	74	0.76%
Interrogative (INTERROGATIVE, 41)	13	0.13%
Discourse Marker (D.M, 43)	9	0.09%
Preposition (P, 42)	5	0.05%
Complex Verb (COMP.V, 44)	2	0.02%

Transitive Verb:

(36) *hem magi'im le-macav she-**menatkim lahem maym**.*
they reach to-situation that-disconnect to.them water.

‘They are reaching a point were **they’re water supply is cut off**.
(22)

Adjective:

(37) *mutar li lehacia od haca’a?*
allowed to.me to suggest another suggestion?
‘Can I make another suggestion?’
(280)

Intransitive Verb:

(38) *im ha-hore lo yaxol leshalem ve-**xozeret lo hora’at keva***
if the-parent not can to pay and-returns to.him standing order
[. . .] ha-irgun yoce nizok.
[. . .] the-organization exits damaged.
‘If the parent can’t pay and **his standing order bounces** [. . .] the organization is damaged.’
(383)

Adverb:

(39) *xaval li leakev et ze.*
pity to.me to delay ACC it.
‘I wouldn’t like to delay it (lit. It would be a pity if I delayed it).’
(14801)

Noun:

(40) *misrad ha-bri’ut shutaf lanu.*
ministry the-health partner to.us.
‘The health ministry is our partner.’
(4,638)

Interrogative:

(41) *lama lanu arba kupot xolim?*
why to.us four health maintenance organizations?
‘Why do we need four health insurance companies?’
(1,528)

Preposition:

(42) *le-oneg li lehishtatef b-a-diyun.*
to-pleasure to.me to participate in-the-discussion.
'It's a pleasure for me to participate in this discussion.'
(3,232)

Discourse Marker:

(43) *yed'u na ha-cofim be-migrashey ha-kaduregel*
should know please the-viewers in-fields the-football
*she-**xalila** **lahem** **linkot** **be-alimut**.*
that-God forbid to.them to use in-violence.
'Let the football audience know that **they mustn't use any kind of violence**.'
(15,514)

Complex Verb:

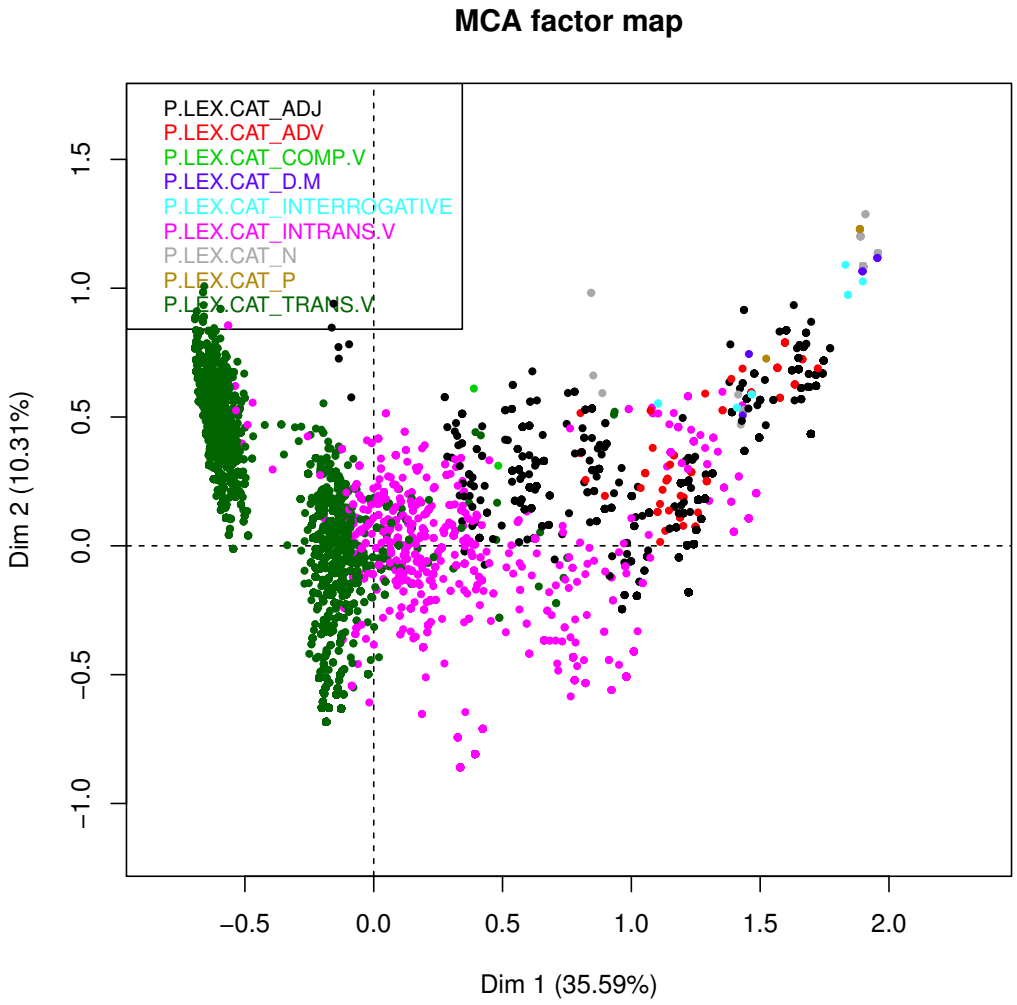
(44) *ze meragesh ve-ze **me'orer** **hashra'a** **lanu** ke-irgun.*
it exciting and-it invokes inspiration to.us as-organization.
'It's exciting and **inspiring for us** as an organization.'
(14,390)

Figures 4.12 and 4.13 present the distribution and correspondences of the lexical categories in the corpus. In most cases, tokens with the same lexical category are adjacent on the MCA map. This fact suggests that the lexical category of the predicate is a good source for between-token similarity. That is, considering previously discussed parameters, Dative constructions may be differentiated not only by their semantics (the semantic type of predicate they involve) or syntax (number, types and order of arguments), but by the lexical category of their predicate as well. The distribution of lexical categories strengthens the affectedness scale proposed earlier, ranging from highly affecting transitive verbs on the left, through the intransitive verbs at the center of the map, to adjectives and adverbs towards the right, with other low frequency non-affecting lexical categories scattered at the right edge.

4.2.2.6 Verbal paradigm (Binyan)

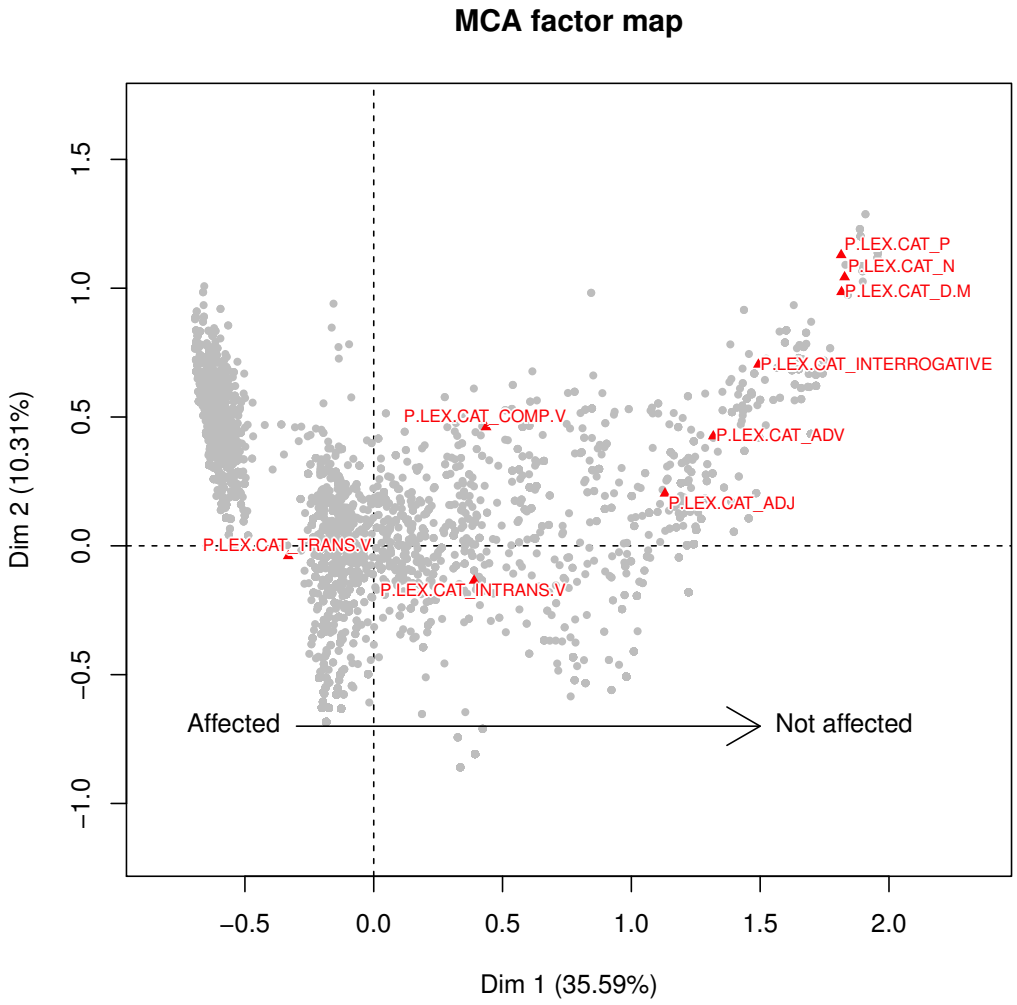
A Hebrew predicate can be classified in another way, on top of its lexical category, namely, its morphology. Each verbal predicate in Hebrew belongs to a morphological paradigm, the *Binyan* system. Looking at Figure 4.14 we can see the distribution of tokens according to their predicate's verbal paradigm. The two transitive

Figure 4.12: Multiple Correspondence Analysis: lexical category



paradigms are hard to tell apart on the MCA map: the *Hifil* paradigm in black and the *Piel* paradigm in grey. These two are grouped together with the multifunctional paradigm *Kal* on the left side of the map. The left part of the map's right side is occupied by the intransitive paradigms *Hufal*, *Pual*, and *Hitpael*, while the far right is filled with dark blue points, representing the IRR paradigm (for irrelevant, i.e., non-verbal predicates). This distribution reflects the lexical categories distribution discussed in section 4.2.2.5; paradigms related to highly affective construals appear on the left, less affective are located at the center of the map, and

Figure 4.13: Multiple Correspondence Analysis: lexical category II



the right edge is reserved for non-affective construals. However, considering the correspondences between the paradigms illustrated in Figure 4.15 one can notice subtle differences between different paradigms that are supposed to be related to the same type of construal. Although both *Piel* and *Hifil* are transitive paradigms, they are located on different positions on the vertical dimension of the Multiple Correspondence Analysis map: the *Piel* at the top half and *Hifil* belongs to the bottom half of the map. That is, assuming that the results of the Multiple Correspondence Analysis represent usage profiles, the two transitive paradigm are

related to different profiles.

A second opposition can be noticed in Figure 4.15 between the intransitive paradigms *Pual* and *Hufal* on one hand, and *Nifal* and *Hitpael* on the other: *Pual* and *Hufal* are located closer to the origin point. That is, they belong to usage patterns similar to the ones characterizing a mild effect on the Dative-marked participant. *Nifal* and *Hitpael*, on the other hand, are located further to the right; they belong to usage profiles related to low (or none) affectedness of the Dative-marked participant.

Figure 4.14: Multiple Correspondence Analysis: verbal paradigm (the Hebrew *Binyan* system)

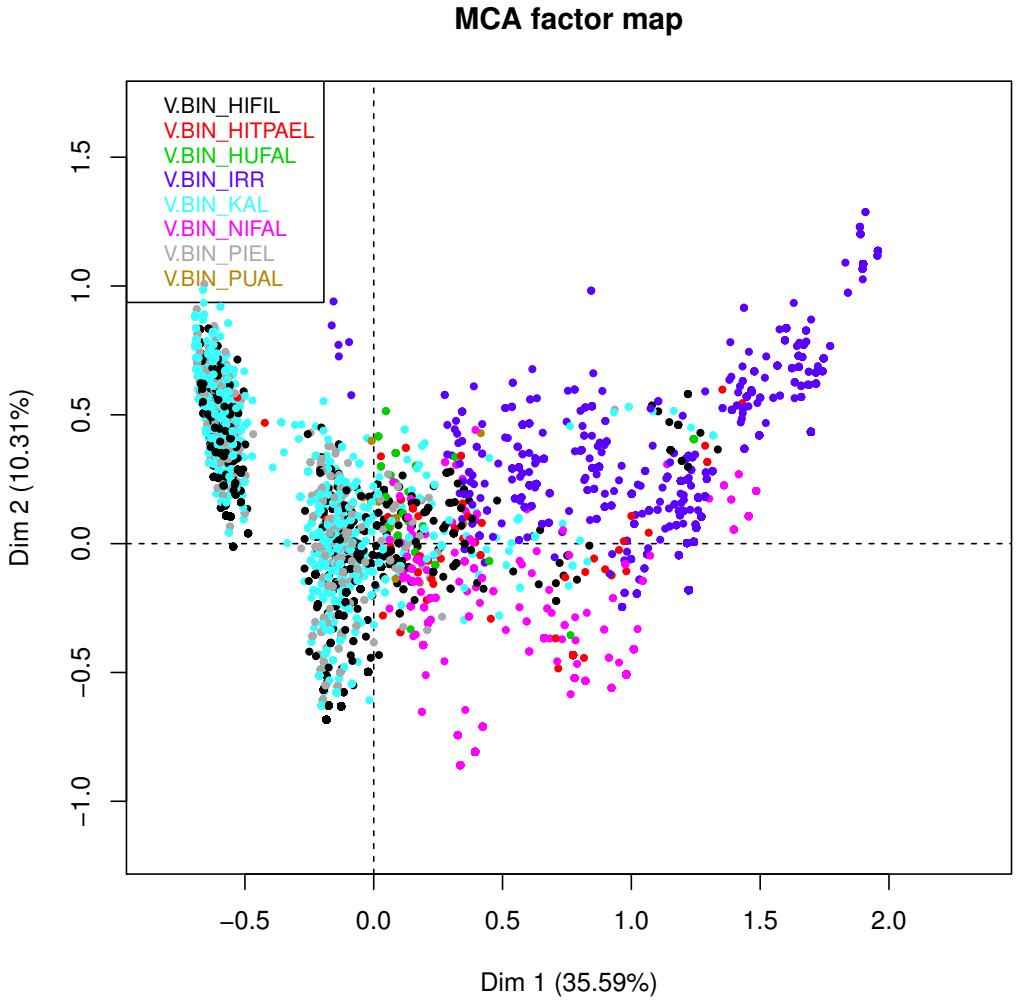
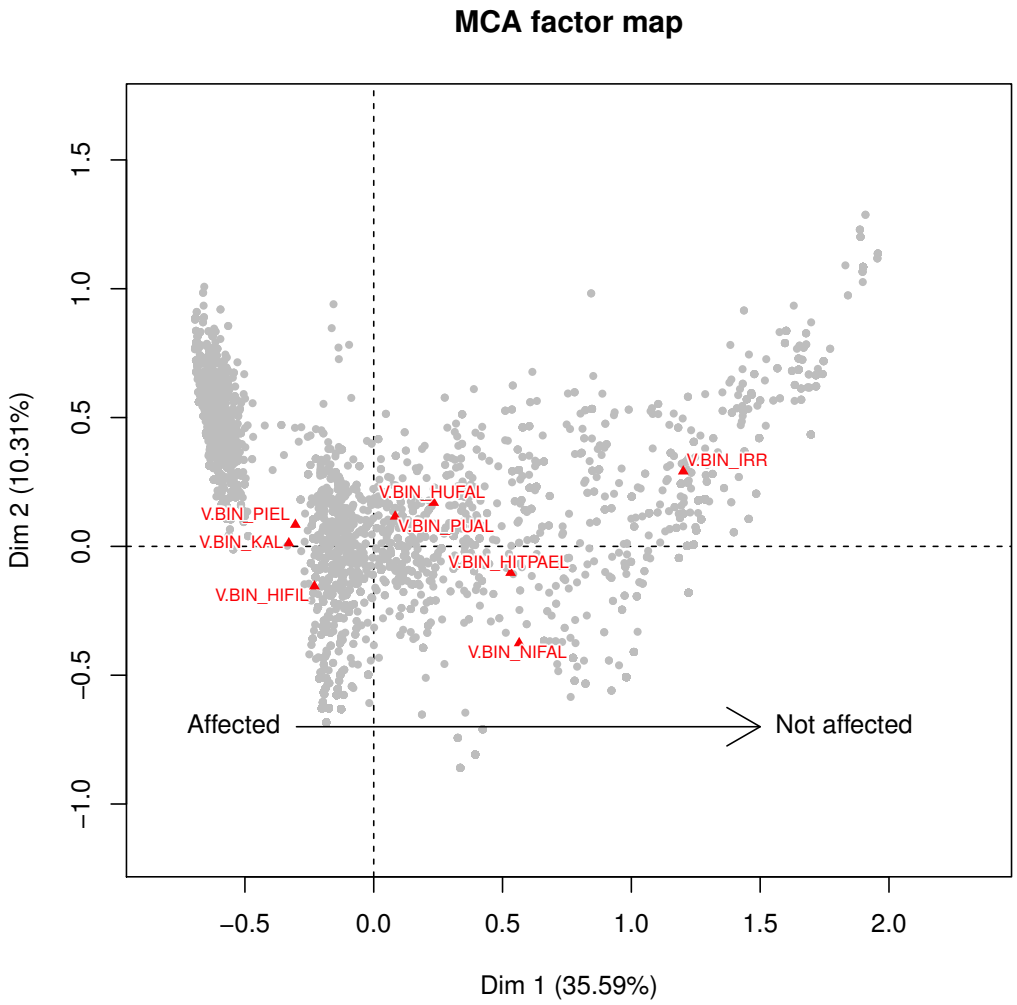


Figure 4.15: Multiple Correspondence Analysis: verbal paradigm (the Hebrew *Binyan* system) correspondences



To sum up the discussion about the morpho-lexical parameters of the Dative clauses, we have seen a three-way division of the map correlated with a level of lexical and morphological transitivity. The distribution of Dative tokens, according to the Multiple Correspondence Analysis, is regulated by lexical and morphological features of the clause's predicate. That is, we can treat the morpho-lexicon as another emergent organizing principle for the Dative category, joining other principles identified above such as the semantic type of the predicate and the syntactic parameters of word order and type and number of arguments in the clause.

4.2.2.7 Mode

The *mode* parameter, coding realis vs. irrealis clauses, was shown earlier to be linked to the vertical dimension. Figure 4.16 shows the manifestation of this link. Here we can see another aspect of the Dative distribution, that of Subjectivity. While the Affectedness Scale proposed earlier is correlated with the horizontal dimension, tokens from either sides of the X axis (above or below it) can be differentiated by Subjectivity parameters. Thus, as we have seen in the discussion about the predicate types, the usage pattern of the tokens placed on the bottom half of the map is related to a construal of a subjective stance, usually by a cognition predicate (and see also Verhagen (2005)).

4.2.2.8 Person of the Dative-marked participant

A support for the partition of the map according to levels of Subjectivity on top of the Affectedness Scale comes from looking at the *person* of the Dative-marked participant. Once again, we can see that the tokens on the bottom half of the map are characterized by first and second person Dative-marked participants, indicating subjective construals, while the top half of the map is related to third person Dative-marked participants, indicating less subjective construal to the extent that the speaker/hearer are not onstage (Langacker, 1990).

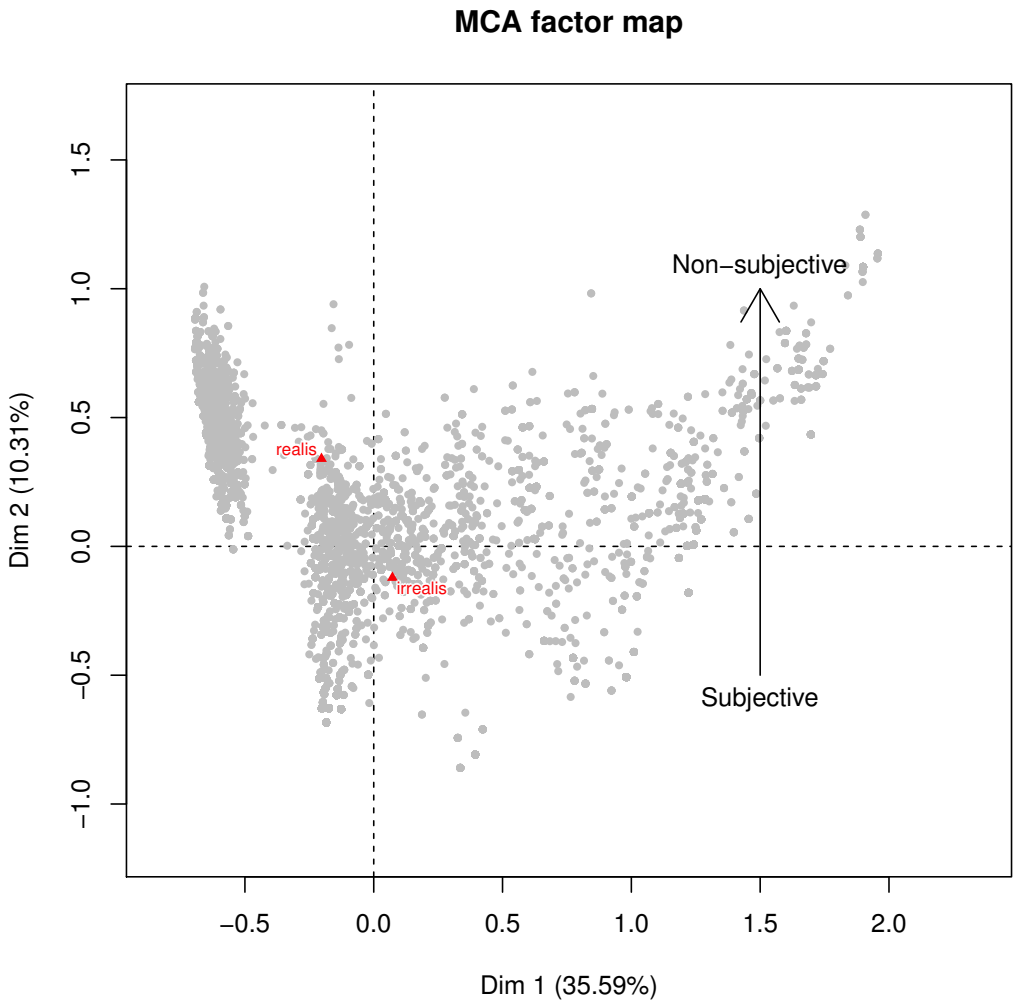
4.2.2.9 Agentivity of the A participant

We turn now to consider two context-based parameters of transitivity: Agency of the A participant and Individuation of the O participant. The Agency parameter adds another layer of interpretation to the divisions of the map we have already seen: an intersection of the Subjectivity and the Affectedness scales, such that the top-left quadrant of the Multiple Correspondence Analysis map is related to tokens with relatively high degree of Agentivity of the A referent, while the bottom half (left and right) is related to low degree of Agentivity. That is, high degree of effect on the Dative-marked participant is related to non-subjective clauses, and these two are related to high degree of Agentivity of the A referent.

4.2.2.10 Individuation of the O participant

The last parameter we consider is the Individuation of the O referent. This parameter is only relevant to three-argument structures with a nominal Direct Object. Thus, it is relevant mainly to the top-left quadrant of the Multiple Correspondence Analysis map. In Figure 4.19 we can see a division within this quadrant such

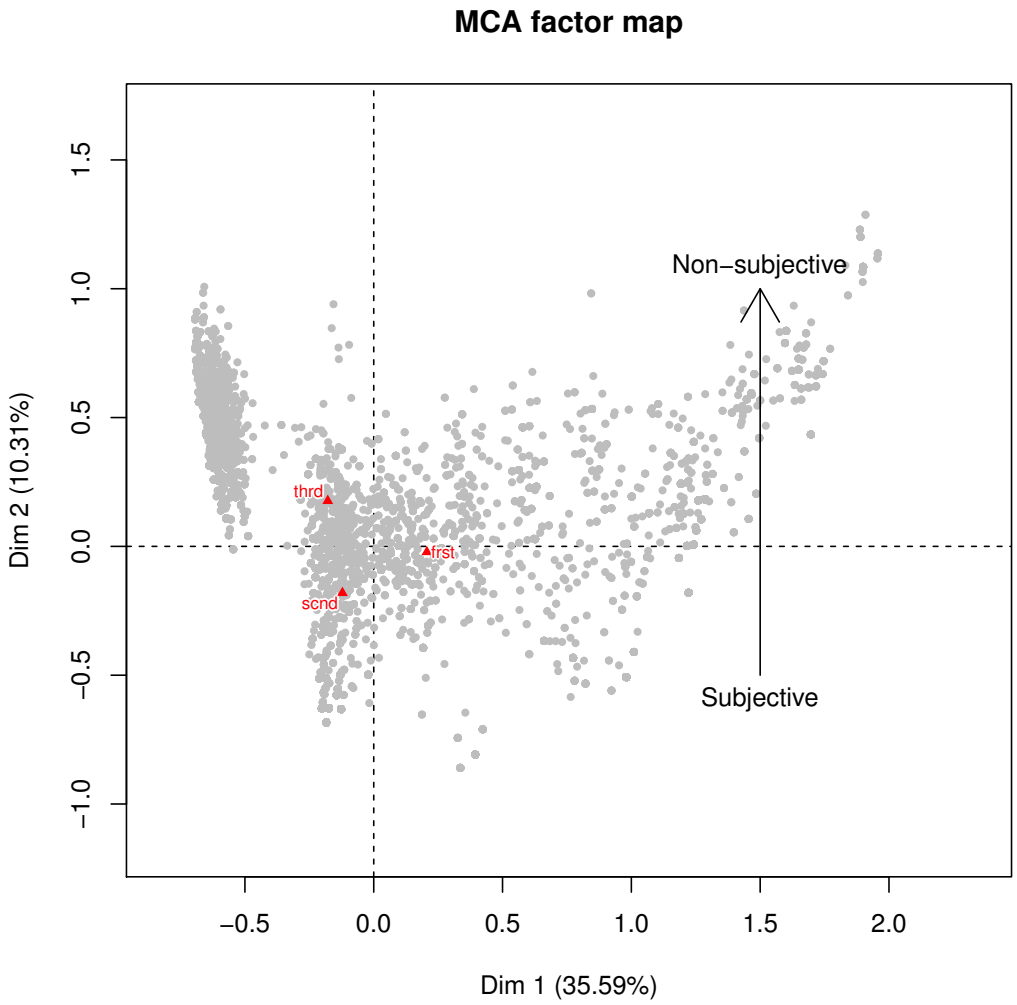
Figure 4.16: Multiple Correspondence Analysis: Mode correspondences with Subjectivity scale



that clauses with high degree of O individuation are located higher than clauses with low degree of O individuation. However, this division is hardly interpretable through the Multiple Correspondence Analysis map.

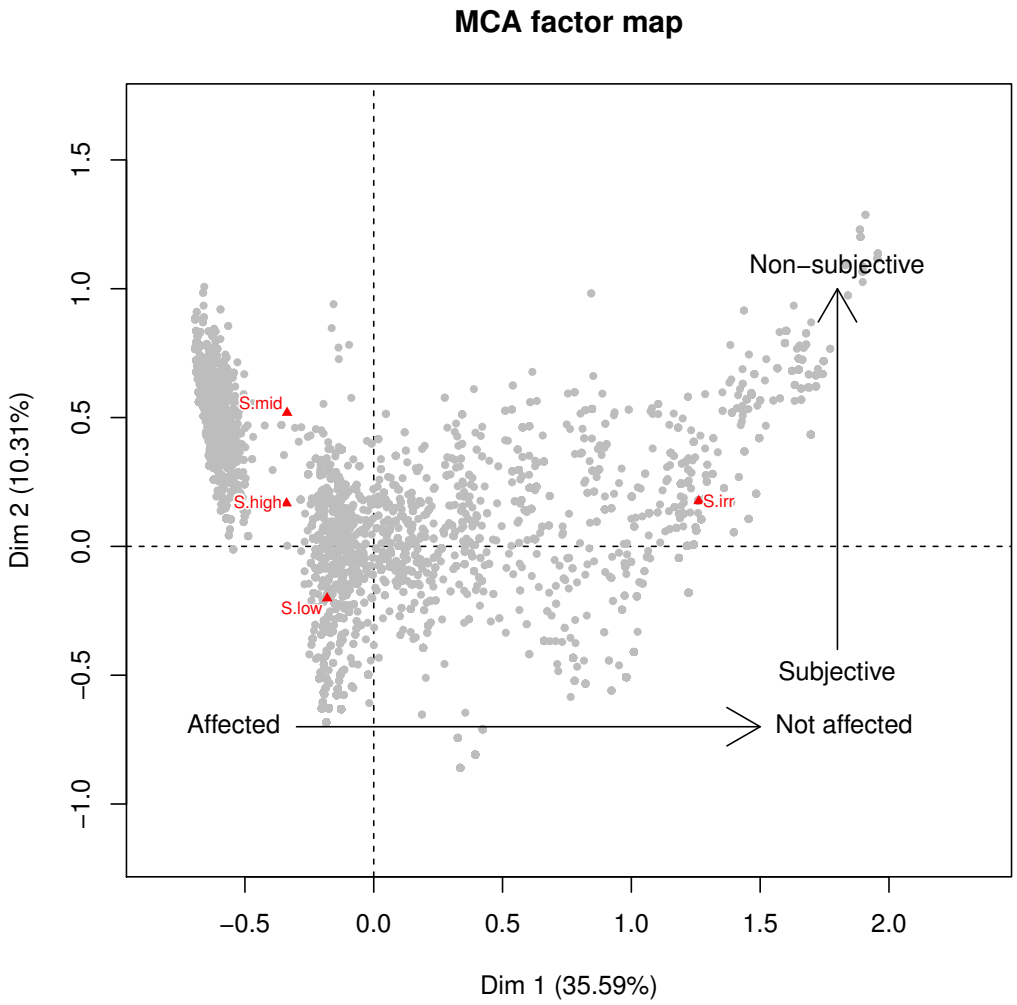
In order to better understand the O individuation parameter we can perform a Correspondence Analysis that shows differences between low and high degree of the O in the context of Dative functions. Figure 4.20 presents the results of this Correspondence Analysis. Here we can see that this parameter is linked to a very

Figure 4.17: Multiple Correspondence Analysis: correspondences between Person of the Dative-marked participant with Subjectivity scale



interesting and illusive issue: the difference between the Possessive Dative and the Affectee. In the Correspondence Analysis map we can see that high degree of O individuation is related to the Possessive Dative function, in the sense that tokens marked for the Possessive Dative tend to be marked for high/mid O individuation. Conversely, tokens with mid-low degree of O individuation tend to be interpreted as Affectee or Recipient. That is, based on a Correspondence Analysis we can conclude that a Possessive Dative interpretation is related to clauses with a higher

Figure 4.18: Multiple Correspondence Analysis: Agency of A with Affectedness and Subjectivity scales



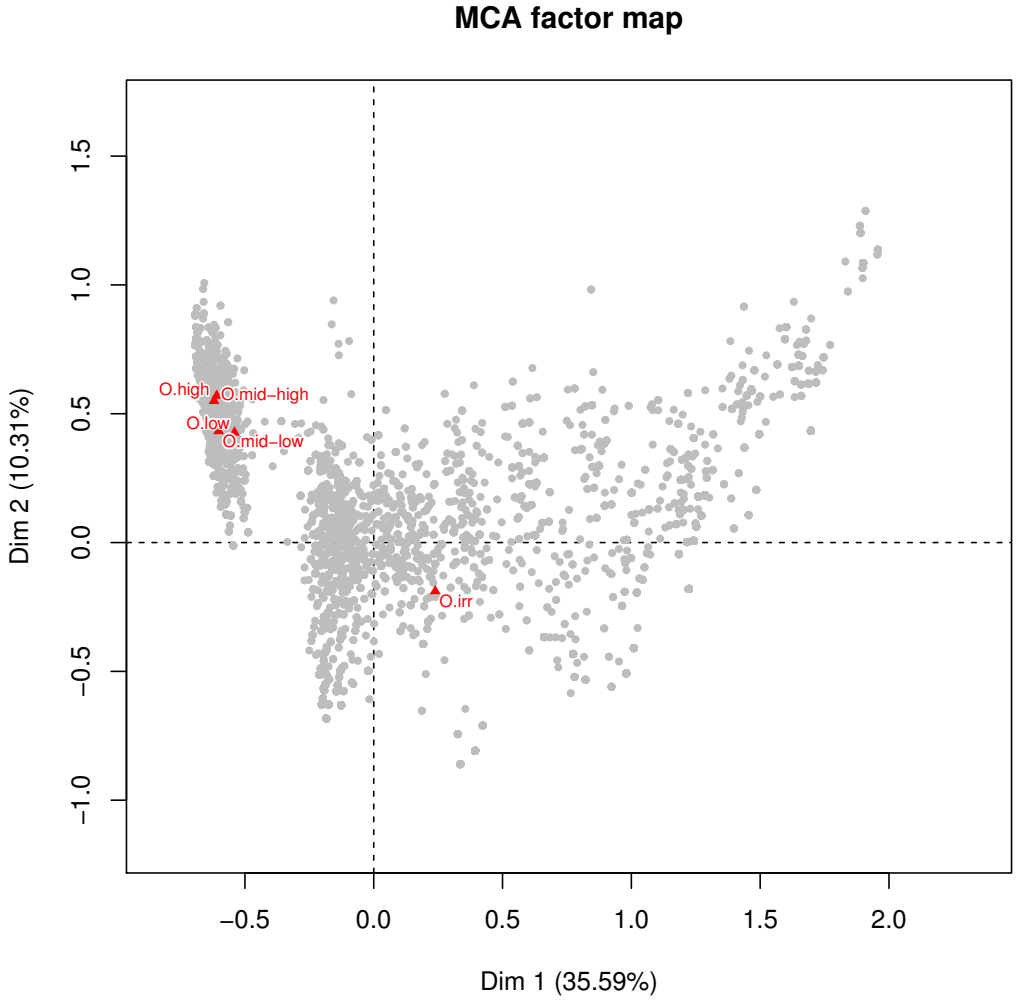
degree of transitivity than an Affectee interpretation. Consider in this connection the following pair of constructed examples:

(45) a. *patxu lo tik.*
open to.him file.
'He's got issued a criminal record.'

b. *patxu lo et ha-tik.*
open to.him ACC the-bag.

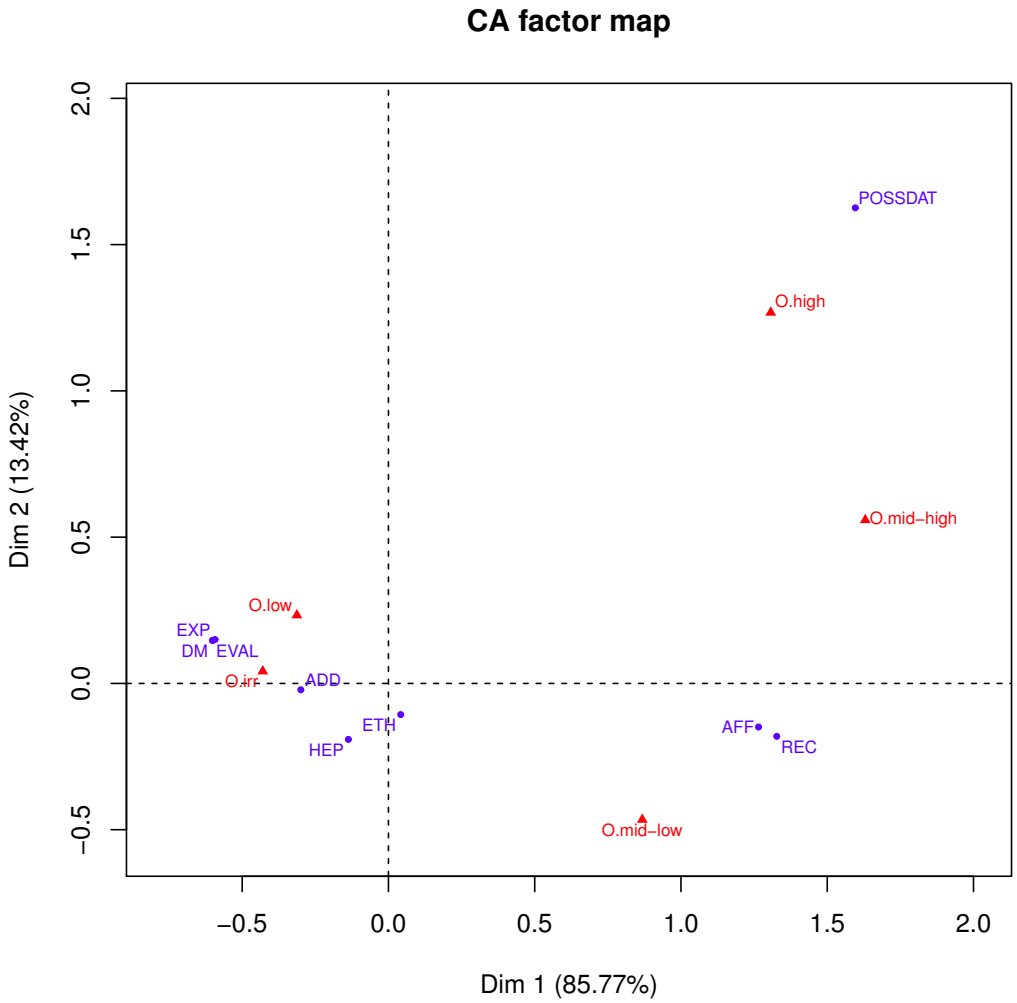
'They've opened his bag.'

Figure 4.19: Multiple Correspondence Analysis: Individuation of O



The only formal difference between the two sentences is the higher degree of individuation of the O referent in the second example. Other than that, formally, they are exactly the same (indeed, *tik* is ambiguous between 'file' and 'bag' but this is not a formal distinction). However, their interpretation is completely different: the sentence with the higher O individuation (45b) receives a Possessive Dative interpretation, while the other is interpreted as an Affectee (45a).

Figure 4.20: Correspondence Analysis: O individuation and Dative function



We can see, then, that there is not one but many candidates for the role of an organizing principle according to which Dative tokens are sorted and stored. In exemplar models of linguistic representation it is assumed that a token's representation may include both individual parameters and an assembly of which as a source for analogies with other tokens, with redundancy of features not interfering. Thus far we have accounted for individual features in isolation, concluding that many of them can function as an organizing principle with a recurrent pattern we have identified as an affectedness scale. In the next section we account for the

Dative interpretation: a subjective coding variable referring to the participant role of the Dative-marked participant as was defined earlier in chapter 3.

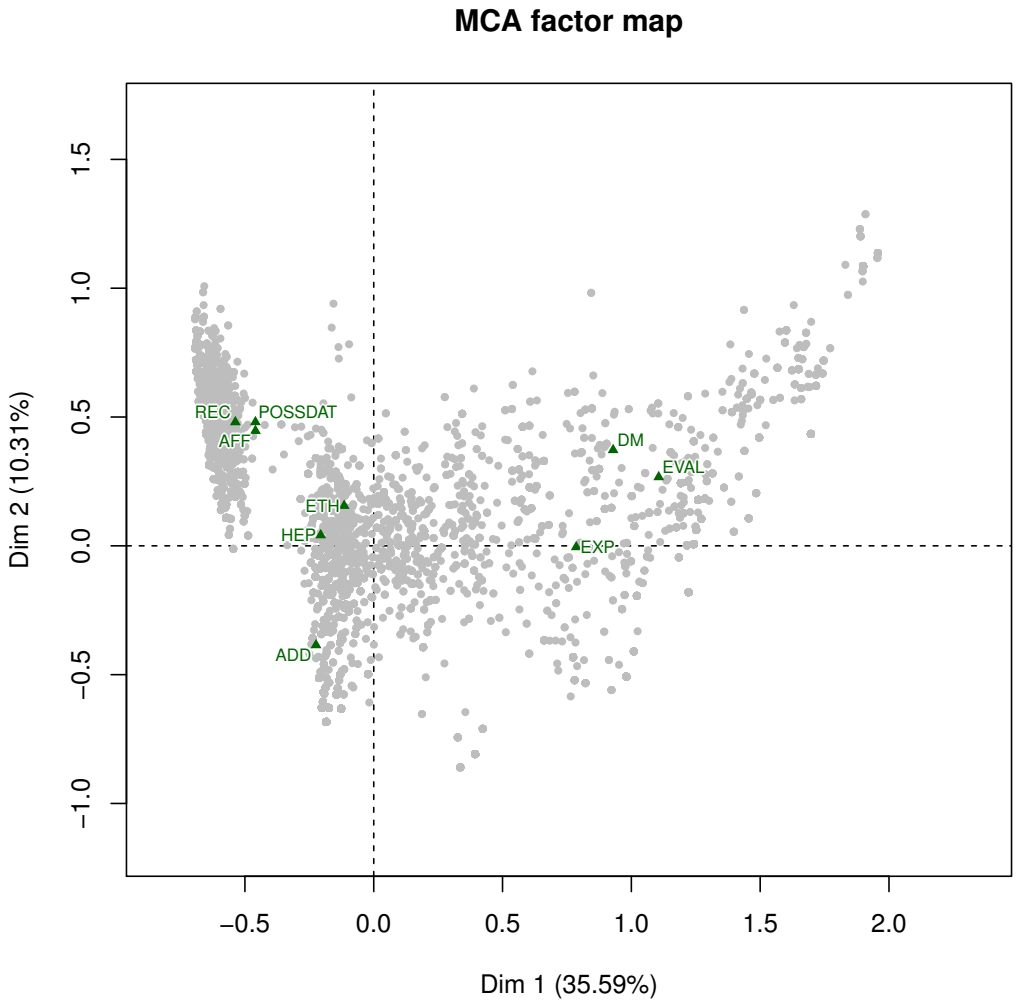
4.2.3 Converging Dative functions

Although the Dative function parameter has not been included in the construction of the map (it is a supplementary variable), it gets coordinates based on the tokens encoded for its particular categories, representing the distance between groups of tokens with different Dative functions. Figure 4.21 presents the locations of the supplementary variable *Dative function*'s categories. The most important point to notice here is that although a token is located on the map regardless of the Dative interpretation it is associated with, we can still see a clear clustering of similar functions; groups of tokens with shared features, similar enough to be adjacent on the map, are associated with a similar Dative function as well. That is, the interpretation of the Dative is shared not only among tokens with similar syntactic structure, as was hypothesized above, but among tokens with multiple types of similarities. The clusters of functions revealed in Figure 4.21 resemble what we have already seen above. Two distant (dissimilar) patterns emerge: the Recipient, the Affectee, and the Possessive Dative pattern together on the top-left quadrant, and the Experiencer and the Evaluative pattern together at the right half. The tokens encoded with an Addressee Dative interpretation have some similarity to the set of Recipient-Affectee-Possessive, but are clearly distinguished from them. The Human Endpoint and the Ethical Dative are similar in some respects to the cluster of Recipient-Affectee-Possessive, and in other respects to the Addressee. The Discourse Marker shows behavioural similarity to the Experiencer-Evaluative set, indicated by its location on the right half of the MCA map.

Figure 4.22 offers another look at the Dative function variable, with tokens coloured according to the Dative interpretation they were coded for. Although there is some mixture of functions (i.e., it is not the case that each function is isolated), the sets of functions discussed above can be easily distinguished: the right half of the map is populated with pink and light blue dots, representing tokens with an Experiencer or Evaluative Dative functions respectively, while the left half of the map is occupied with black, dark-green, brown, and red points, representing the Addressee, the Recipient, the Possessive Dative, and the Affectee, respectively. The Human Endpoint function, represented by grey dots, is scattered around the origin with some tendency towards the top-left quadrant.

Summing up the Multiple Correspondence Analysis of the present corpus, we can put forward the following conclusions. First, other than the predicate itself, which functions as a clue for the particular Dative interpretation (see chapter 1), no

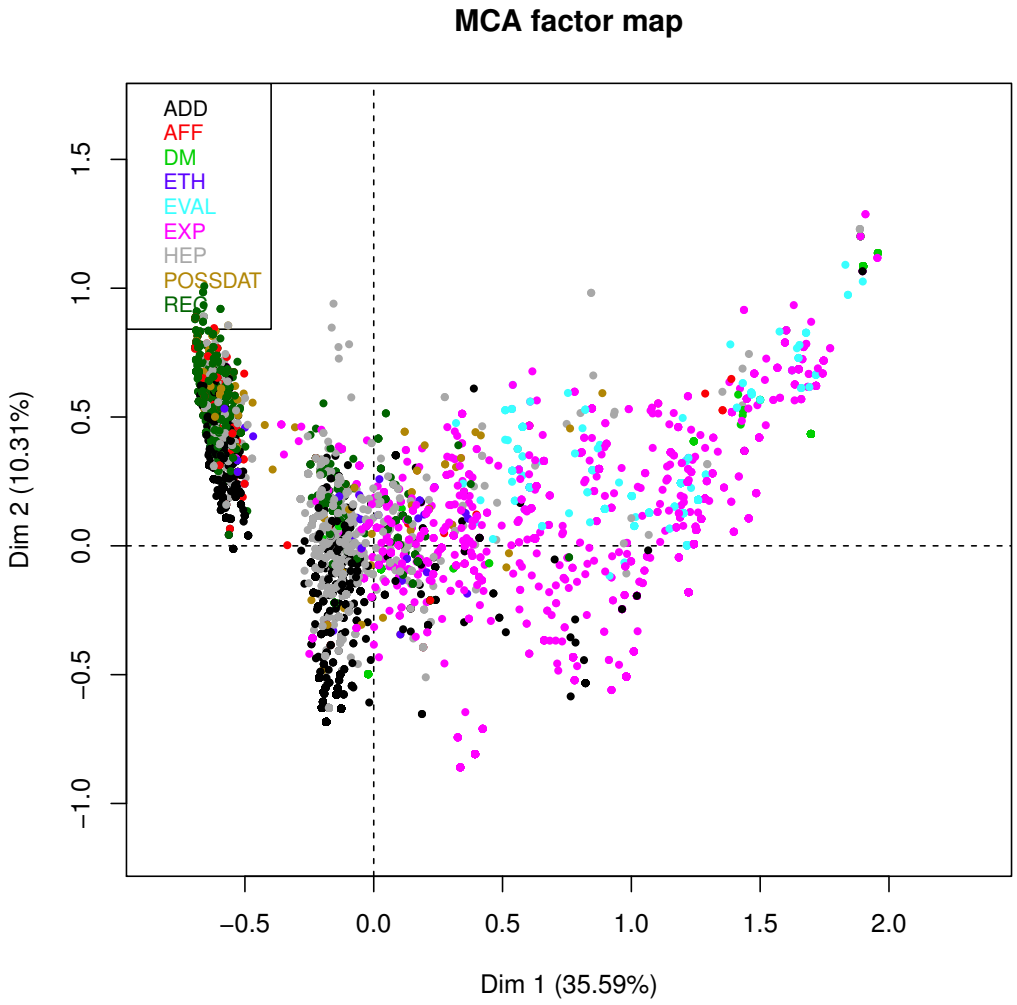
Figure 4.21: Multiple Correspondence Analysis: Dative functions



abstract, single parameter was found to be correlated with a single Dative participant role. Rather, only by considering multiple parameters, from multiple sources of information, can we identify a correlation between form and function (form being the set of parameters, function being the Dative interpretation).

Moreover, we have seen that Dative functions are not isolated. Rather than observing a dichotomous distribution of Dative functions, we have seen a gradience of functions on the MCA map. We *can*, however, observe a dichotomy between *sets* of functions. That is, the second conclusion we can draw from the Multiple

Figure 4.22: Multiple Correspondence Analysis: Dative functions

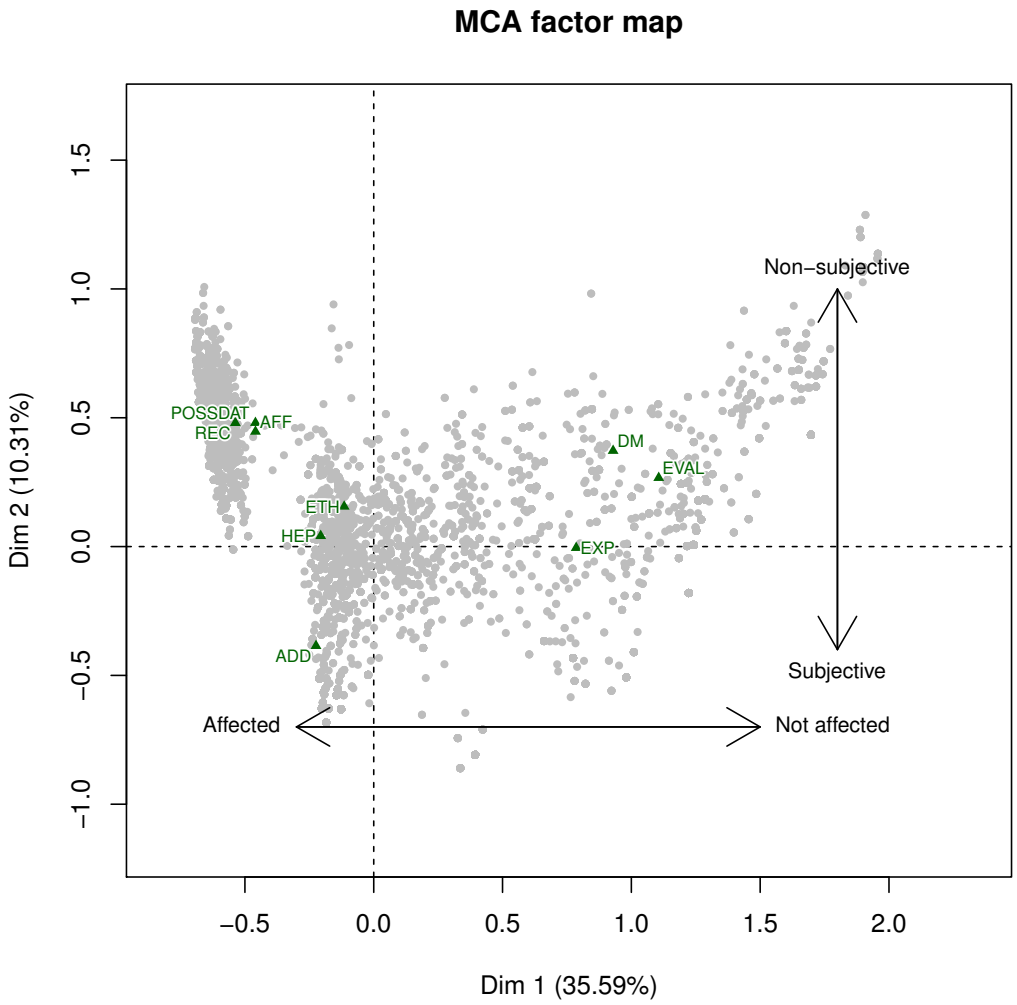


Correspondence Analysis is that in the context of usage patterns we can successfully differentiate only three functions encoded by the Dative in Hebrew, which converge together sets of predicate-specific participant roles: (i) a reference point of an affective situation, (ii) a human endpoint, and (iii) a reference point of a non-affective situation.

The distribution of these sets of functions is not arbitrary. We have found several parameters which constitute a source for a comparison between tokens, attaching groups of tokens together. We have seen that although the Dative func-

tion was not taken into account in the construction of the map, tokens grouped together according to a particular set of parameters tend to share a Dative interpretation. The relevant feature unifying this set of parameters is the degree of effect exerted upon the Dative-marked participant. That is, affectedness plays a major role in structuring the Hebrew Dative category, but not in the sense suggested by Berman (1982), as a general, categorical, Affectee marker. Rather, tokens of Dative clauses are interpreted as similar or different according to the *degree of effect* exerted upon the Dative-marked participant, a degree that varies from no effect at all to a complete effect that changes the referent's state. The third conclusion drawn from the Multiple Correspondence Analysis, then, is that the organizing principle of the Hebrew Dative category is an Affectedness Scale. That is, rather than being a function of a particular syntax or a specific semantic or pragmatic context, the different Dative interpretations are located on a scale of affectedness which is composed of a combination of usage patterns. A token of use is interpreted according to the values it receives for each of the parameters discussed above. A particular configuration of these values is a usage pattern that places the token on a specific point on the Affectedness Scale, over which the Dative functions are spread. This scale can be seen in the MCA maps presented above; tokens with highly affected Dative-marked participant on the left, partially affected Dative-marked participant in the middle, and participants mildly affected (or nat at all) located on the right. Figure 4.23 illustrates the Affectedness Scale with reference to the fine-grained, predicate-specific, Dative functions. In the next chapter we will see that this partition of the data according to an Affectedness Scale can be backed up by cluster analysis.

Figure 4.23: Affectedness Scale



5
Cluster analysis

In the previous chapter a corpus of Hebrew Dative clauses has been analysed using Multiple Correspondence Analysis, thus revealing hidden patterns in the data. We have indicated that the data can be divided into several clusters, each of which corresponds to a set of converged Dative functions. In the present chapter we perform a clustering analysis in order to show that the visually detected clusters proposed in Chapter 4 indeed group together similar tokens. We show that each group of similar tokens represents a discourse pattern: a combination of particular parameters from many sources of information – both linguistic and extra-linguistic. Each such discourse pattern correlates with a single type of construal of events. This correlation is defined as a Discourse Profile Construction. Discourse Profile Constructions are Argument Structure Constructions in a broad sense. They are conventional form-function pairings, combining a multifactorial form with a single construal of the world. Moreover, each Discourse Profile Construction unifies different Dative participant role according to the level of affectedness of the Dative-marked participant.

We will see that the clustering process yields five clusters. Each of the five clusters in the data is both a cluster of tokens, and a cluster of lexical, morpho-syntactic, semantic, and discursive features. Each such cluster is paired with a different construal. In the present context of accounting for the Hebrew Dative, different construals mean different sets of converged Dative interpretations.

Each of these clusters is differentiated from the others based on transitivity and

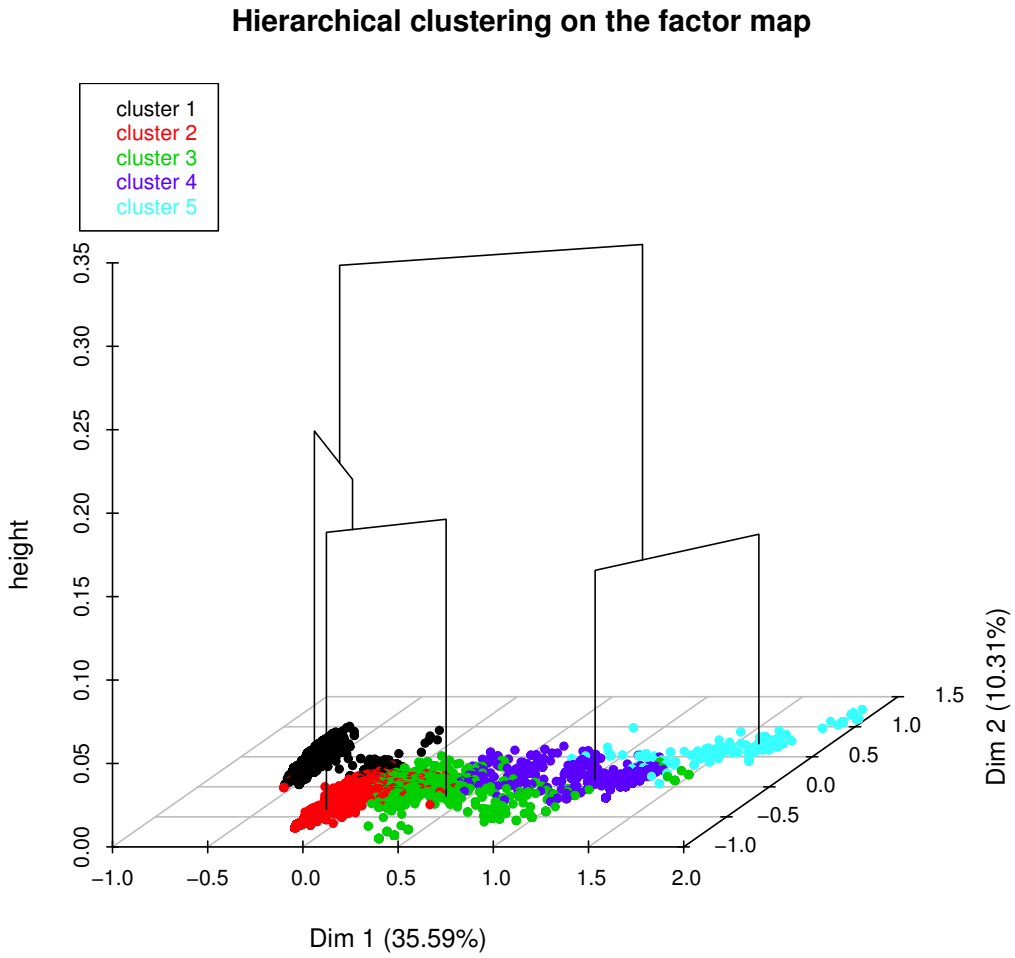
subjectivity parameters, combined together to form an Affectedness Scale. This scale is what guides both the interpretation of the clause and the categorization of tokens into clusters. For instance, a clause with a high degree of transitivity and a low degree of subjectivity is similar to the tokens in Cluster One. It will then be categorized with them, and will be interpreted accordingly as related to a high degree of effect exerted on the Dative-marked participant. Such a high degree of effect is a converged function, composed of the participant roles Recipient, Affectee, and Possessive Dative. That is, the speaker/hearer (or the linguist, for that matter) may not be able to decide whether the right interpretation is a Possessive Dative, for example, or an Affectee. Rather, they interpret the clause according to the level of affectedness of the Dative-marked participant, as High, for example, and the particular participant role is dictated by particular parameters (type of verb, for instance, or degree of individuation of the O referent).

5.1 Introduction to clustering

Exploratory statistics provides us with tools for uncovering structure in the data in a bottom up manner. So far we have used Multiple Correspondence Analysis in order to find the parameters that reflect the distribution of the data, concluding that only an assembly of parameters can function as an organizing principle for the Dative category. Intuitively, we have visually noticed that the data is composed of clusters of similar tokens of use, each cluster corresponding to a set of Dative functions that share a usage pattern. However, so far the identification of the said clusters was not backed up by anything rather than the apparent adjacency of the dots on the Multiple Correspondence Analysis map. In order to find out what clusters emerge from the data in a bottom-up, statistically guided approach, I use a complementary method to the MCA, namely, Hierarchical Clustering on Principal Components (HCPC, introduced in section 2.3).

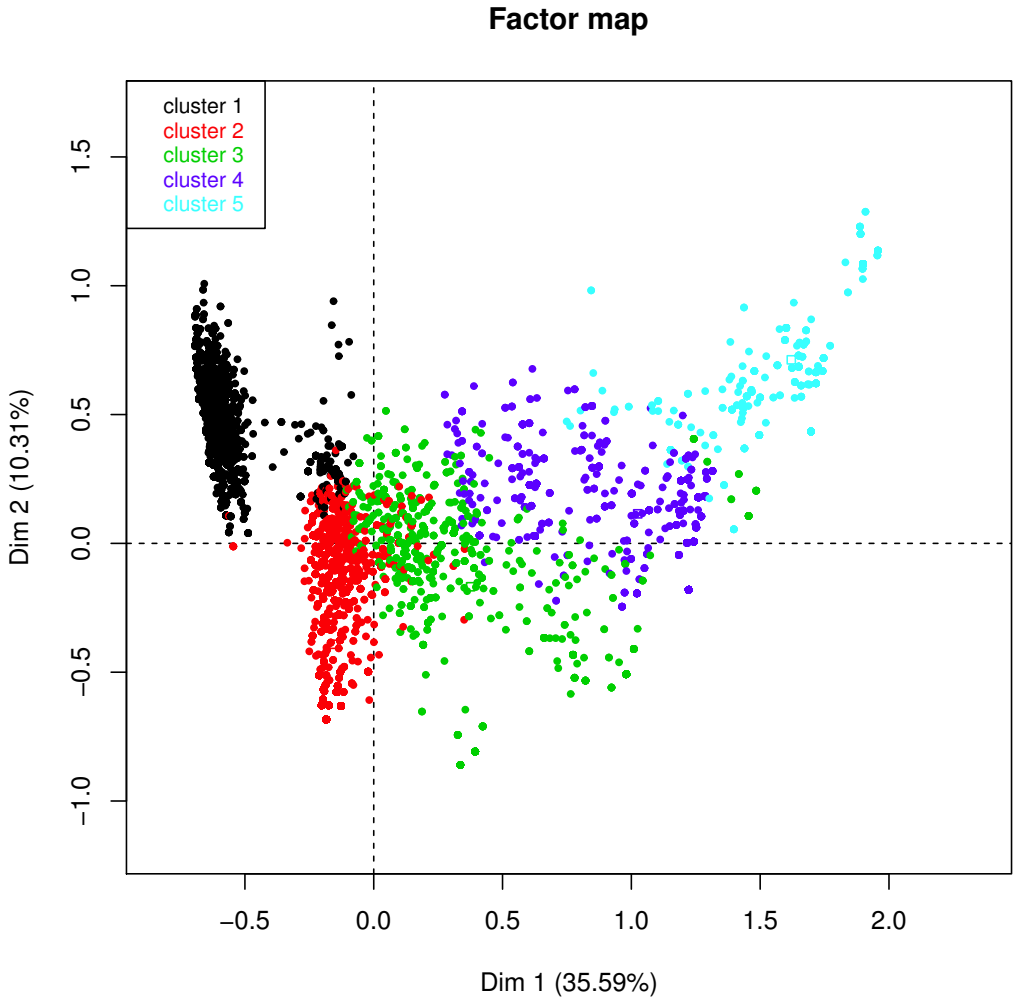
Coming to analyse the results of the Hierarchical Clustering on Principal Components, we first need to understand how we describe and interpret the clusters. For example, consider Cluster Four of the present corpus. The fourth cluster identified by the clustering analysis consists of 1,141 tokens, which are 11.7% of the corpus. As can be seen in Figures 5.1 and 5.2, the tokens belonging to Cluster Four are located on the right half of the MCA map. Recall that in the previous chapter we have concluded that the right half of the map corresponds to low (or no) effect on the Dative-marked participant. Interpreting the results of the cluster analysis we can learn more about these tokens, their similarities and dissimilarities, and their grouping.

Figure 5.1: Hierarchical Clustering on Principal Components: tree diagram



Cluster Four is a non-verbal cluster; 92.5% of its tokens have an adjectival predicate, 3.5% adverbial, 2.7% Intransitive Verb, 1% Transitive Verb, and 0.17% Complex Verb. Some of these numbers are very small, and by themselves may not reveal much about the cluster's tokens. However, it is not the only way to account for frequency of categories in the data. Another perspective would be to count how many Adjectival tokens, for instance, belong to Cluster Four compared to their percentage in other clusters. Through this perspective we can see that 81.6% of the tokens with Adjectival predicates belong to Cluster Four. That is, a

Figure 5.2: Hierarchical Clustering on Principal Components: two-dimensional map



description of a cluster X according to its variables should take into account the percentage of tokens with a category Y that belong to cluster X, the percentage of the tokens in cluster X that manifest the category Y, and the global percentage of the tokens manifesting category Y in the corpus (and see in this connection the debate regarding collostructional analysis, Stefanowitsch and Gries (2003); Bybee (2010)). Each category can then be assessed with respect to the strength of its link to the cluster.

Figure 5.3 on the next page provides a fragment of Cluster Four's description according to the coded variables in the corpus. The first column lists the categories of every variable in the corpus. The second column (Cla/Mod) presents the percentage of tokens manifesting each category that belong to the cluster. The third column (Mod/Cla) indicates the percent of the tokens in the cluster that possess each category. The fourth column (Global) lists the frequency of each category in the corpus. The fifth column (p.value) indicates whether a category is linked to the cluster or not. A p.value of less than five percent suggests a link. The last column (v.test) is related to the representation of the category in the cluster. If the v.test value is positive, the category is over-represented; if it is negative, the category is under-represented. Consider the fourth row of the table in Figure 5.3 for example. The second column (Cla/Mod) indicates that 81.67% of the tokens coded for 'Predicate Lexical Category = Adjective' belong to Cluster Four. The third column shows us that 92.5% of the tokens in Cluster Four manifest the category *Adjective*. In the fourth column we see that the tokens manifesting the category *Adjective* constitute only 13.3% of the corpus. That is, we can see that while the category *Adjective* is relatively rare in the corpus (occurring in only 13.3% of the tokens), it constitutes the majority of Cluster Four. And indeed, the last column indicates that the v.test value for the representation of the category in the cluster is positive infinite, meaning that it is over-represented. As a comparison, the category *Transitive Verb* on the second-to-last row has a negative infinite value in Figure 5.3. That is, Cluster Four is strongly linked to the *Adjective* category, while it has a very weak link to the category *Transitive Verb*.

It is important to note that it is not the case that a cluster consists of only the categories it is linked to. Rather, there may be tokens belonging to a certain cluster that manifest some category which is not characteristics of this particular cluster. The cluster's description provided by the R package, however, solves this problem. For example, the Dative-marked participant of 0.3% of the tokens in Cluster Four is interpreted as an Addressee. This interpretation, we concluded earlier, corresponds to a higher degree of transitivity than the degree associated with Cluster Four's tokens. Nevertheless, we can still say that Cluster Four is not characterized as having an Addressee interpretation since the link between the cluster and the category is very weak, compared, for instance, to the link between the cluster and other categories of the same variable, namely the Experiencer or the Evaluative Reference Point (see Figure 5.3). The interpretation of tokens belonging to Cluster Four as having an Addressee Dative function will be within the limits of a low transitivity context, receiving nuances it would not have had it occurred in a high transitivity context. We will address this issue in later sections.

A cluster can be described, among other things, according to the categories it is

Figure 5.3: Description of Cluster Four by the categories in the corpus– a fragment

$category$‘4‘

	Cla/Mod	Mod/Cla	Global	p.value	v.test
V.BIN=V.BIN_IRR	72.5231176	96.23137599	15.61790798	0.000000e+00	Inf
S.AGE.INDIV=S.irr	53.5156250	72.04206836	15.84485249	0.000000e+00	Inf
PRED.FIRST=predFirst	48.7580994	79.14110429	19.10460078	0.000000e+00	Inf
P.LEX.CAT=P.LEX.CAT_ADJ	81.6705336	92.55039439	13.33814731	0.000000e+00	Inf
NO.OF.ARGUMENTS=two	36.9016818	100.00000000	31.89601816	0.000000e+00	Inf
DAT.FUNC=EXP	41.3150685	66.08238387	18.82607799	4.764769e-320	38.246450
DAT.FUNC=EVAL	86.8292683	31.20070114	4.22942026	1.321337e-289	36.368893
P.TYPE=P.TYPE_PROPERTY	72.7436823	35.31989483	5.71487518	1.015726e-272	35.284487
P.TYPE=P.TYPE_VALUE	76.3796909	30.32427695	4.67299360	6.646137e-243	33.283701
O.INDIV=O.irr	16.4243558	100.00000000	71.66288426	1.434227e-178	28.492906
O.DEF=O.def.irr	16.4243558	100.00000000	71.66288426	1.434227e-178	28.492906
DAT.PRSN=frst	20.7707222	75.10955302	42.56240974	8.416928e-125	23.761199
NO.D=NO.D_V	33.6322870	32.86590710	11.50195998	3.391922e-97	20.921783
P.TYPE=P.TYPE_SD	33.9517625	16.03856266	5.56014029	9.809762e-45	14.032872
P.LEX.CAT=P.LEX.CAT_ADV	54.0540541	3.50569676	0.76335878	1.012120e-18	8.833763
O.INDIV=O.mid-high	0.0000000	0.00000000	8.44852486	2.128761e-47	-14.461256
NO.D=NO.D_NP	6.9052103	28.92199825	49.29853518	5.472189e-50	-14.866084
V.BIN=V.BIN_PIEL	0.0000000	0.00000000	9.04683309	7.076433e-51	-15.002445
DAT.PRSN=scnd	3.5214086	7.71253287	25.77883227	6.398153e-61	-16.466385
DAT.FUNC=HEP	0.9977827	1.57756354	18.60944914	3.260260e-81	-19.086649
DAT.FUNC=REC	0.1218027	0.17528484	16.93831236	1.246898e-94	-20.638172
S.AGE.INDIV=S.high	0.7522332	1.40227870	21.94140706	1.310060e-104	-21.720618
O.INDIV=O.mid-low	0.0000000	0.00000000	18.70229008	1.110626e-110	-22.353732
P.TYPE=P.TYPE_PA	1.8329939	4.73269062	30.38993192	1.756458e-116	-22.942358
V.BIN=V.BIN_HIFIL	0.5020921	1.05170903	24.65442542	1.857741e-128	-24.112292
O.DEF=O.def.no	0.0000000	0.00000000	21.41530844	7.126340e-129	-24.151928
S.AGE.INDIV=S.low	4.2577675	19.45661700	53.78584692	1.632538e-141	-25.326475
P.TYPE=P.TYPE_PB	1.8473241	5.95968449	37.97194141	2.331473e-155	-26.553637
DAT.FUNC=ADD	0.3050641	0.87642419	33.81473076	7.597510e-200	-30.161424
V.BIN=V.BIN_KAL	0.6463969	2.36634531	43.08850836	4.686406e-251	-33.842453
PRED.FIRST=predNotFirst	3.0349401	20.85889571	80.89539922	0.000000e+00	-Inf
P.LEX.CAT=P.LEX.CAT_TRANS.V	0.1729356	1.05170903	71.58035898	0.000000e+00	-Inf
NO.OF.ARGUMENTS=three	0.0000000	0.00000000	63.42067258	0.000000e+00	-Inf

linked to (as seen above and exemplified in Figure 5.3). In the present case, such a description reveals a particular usage profile of a particular type of Dative clauses in Hebrew, which conventionally combines together linguistic and extra-linguistic features and has a constant function. In other words, the cluster is a Construction in the Construction Grammar sense: a conventional pairing of form and function. However, while constructions are usually defined based on morpho-syntactic or semantic-pragmatic parameters, the present account defines constructions on a wide, multifactorial basis. I therefore call this type of construction a *Discourse Profile Construction*. This concept is an extension of Argument Structure Constructions. Argument Structure Constructions constitute the syntactic-semantic basis for simple clauses and simple event structures in the language (Goldberg, 1995). The Discourse Profile Construction, in this respect, extends the range of phenomena captured by the construction, and provides the basis for the usage conditions conventionally linked to such event structures. Thus, a particular event, or more specifically, a particular construal of a partial/mental effect in the case of the Hebrew Dative, is conveyed through a particular usage pattern; a particular Discourse Profile Construction. In the following sections I analyse the five clusters identified in the corpus as four Discourse Profile Constructions: Cluster One is the Extended Transitive Discourse Profile Construction, Cluster Two is the Human Endpoint Discourse Profile Construction, Cluster Three is the Extended Intransitive Discourse Profile Construction, and Cluster Four is the Evaluative Reference Point Discourse Profile Construction. Cluster Five is a small cluster that cannot receive a homogeneous treatment; we will address this issue in section 5.6

5.2 Cluster One: Extended transitive Discourse Profile Construction

The first cluster of tokens in the corpus gathers together 2,870 Dative clauses that show a similar usage pattern. This cluster of tokens is linked to Transitive predicates from the Primary A type (verbs with concrete participants), mostly in the KAL verb paradigm (*binyan*), but also in the PIEL paradigm. Syntactically, Cluster One is linked to a three-argument structure with a nominal A, a nominal O, and a Dative-marked participant in the second person, functioning as what a traditional participant role analysis would define a Recipient, an Affectee, and a Possessive Dative (see Chapter 3). Cluster One is also linked to a realis mode of clauses, with an intermediate-high degree of Agency of the A referent, and an Individuated O participant. These characteristics of Cluster One correspond to a high degree of affectedness, and its tokens are located accordingly on the top-left

quadrant of the Multiple Correspondence Analysis map.

Besides a description of a cluster according to the categories it is linked to, each cluster has a limited number of central and unique exemplars according to which it can be described. A central exemplar is a token that shares the highest number of features with other tokens that belong to the same cluster. A unique exemplar, on the other hand, is defined as such since it shares the smallest number of features with tokens that belong to other clusters. The following are some of Cluster One's central and unique exemplars:

(1) Central exemplars:

a. *xiyavti et ha-sar [. . .]* ***lehacig lanu ma ha-misrad asa****.*
I.forced ACC the-minister [. . .] to.present to.us what the-office did.
'I forced the minister to show us what his office was doing.'
(189)

b. *ani mevakeshet mi-necigey [. . .]* ***lehavi lanu netunim*** *yoter meduyakim.*
I ask from-the.representatives.of [. . .] to.bring to.us data more accurate.
'I'm asking the representatives of [. . .] to give us more accurate data.'
(2,223)

c. *hem* ***ya'aviru lanu takanot*** *ad ha-rishon be-november.*
they will.pass to.us regulations until the-first in-November.
'They will get some regulations through for us by November first.'
(4,308)

Looking at (1), we can observe a realization of the usage pattern presented above as strongly linked to Cluster One (and see Appendix C for a detailed list of the strengths of links between each category and cluster in the data). In the central exemplars we see Primary A type of predicates – verbs with concrete participants as their semantic roles – embedded in a three-argument structure with a highly agentive A participant (e.g., the minister or the representatives), and a nominal O. While these clauses share some of their features with clauses from other clusters, their overall usage pattern is enough for them to be judged as more similar to tokens from Cluster One than to tokens from other clusters, whether some particular feature is unique to Cluster One or not. Cluster One's unique exemplar, the one that shares the smallest number of features with tokens from other clusters, is the following:

(2) ***atem*** *lo* ***notnim lo et ha-kelim*** *le-hitmodedut im ha-beaya*
you not give to.him ACC the-tools to-cope with the-problem

ha-zot.
the-this.
'You don't give him the right tools for coping with this problem.' (109)

The exemplar in (2) presents an Accusative marked, relatively individuated O participant, a Primary A, a Transitive predicate in the KAL paradigm (*natan* 'give'), and an A participant with an intermediate-high degree of Agentivity. These features, as we have seen above, mostly characterize tokens from Cluster One, and thus are unique to this particular cluster.

The first Discourse Profile Construction is represented in the data by the tokens and features of Cluster One: it is related to high transitivity parameters. In the present section we explore the actual tokens of use belonging to Cluster One as manifestations of an extension of an affective situation Discourse Profile Construction. However, before discussing the Discourse Profile Construction itself, we need to address the issue of affective situations and the difference between core and extended construals.

5.2.1 Affective situations

A prototypical affective situation is a situation in which a human entity acts upon, or creates, a physical object. It is affective to the extent that the object of the action, its Endpoint, is affected by the action of the Initiator. An affective situation can be construed by profiling the INITIATOR and the ENDPOINT of the event (Croft, 1991, 2012), mapping them to A and O syntactic relations (Dixon, 2005). Consider, for example, the following (partly constructed) examples:

(3) a. *shalosh peamim* ***hem*** *sagru* ***et*** ***ha-yecu***.
three times **they** closed ACC **the-export**.
'Three times they've shut down the export.'
(∼13707)

b. *bi-zman Oferet-Yecuka hociu* ***rabanim ishurim***
in-time Oferet-Yecuka issued **Rabbis certificates**
she-medubar be-pikuax-nefesh.
that it's a unique situation.
'During Oferet Yecuka (= a military operation) some Rabbis issued certificates that it is a unique life or death situation.'
(∼3305)

The speakers in (3a)-(3b) profile the INITIATOR and the ENDPOINT participants (marked with a boldfaced font), thus construing an affective situation. It is a

construal of an Initiator either acting upon an object (3a), or creating one (3b). In both cases, the core event is completely represented. However, languages provide a tool for extending the construal of events, referring to a non-core participant and relating it to the event. Hebrew (among other languages, see chapter 2) uses the Dative form to refer to such an entity bearing such a semantic role.

The set of Hebrew Dative semantic roles constituting the class of extensions of an affective situation is composed of the Recipient, the Affectee, and the Possessive Dative. Common to these roles is the place they take in the construal of events, in that they contribute non-core information, rather than the core meaning of the verb (Levin, 1999). These roles are characterized by less than central involvement in, and control over, the event (Lehmann, 2006). Consider the following attested examples, the originals inspiring the versions in (3):

(4) a. *shalosh peamim* ***hem sagru lanu et ha-yecu***.
three times **they closed to-us OM the-export**.
'Three times they've shut down the export on us / they've shut down our export.'
(13707)

b. *bi-zman Oferet-Yecuka* ***hociu lanu rabanim ishurim***
in-time Oferet-Yecuka **issued for us Rabbis certificates**
she-medubar be-pikuax-nefesh.
that it's a unique situation.
'At the time of Oferet Yecuka (= a military operation) some Rabbis provided us with a certificate that it is a unique situation from the Jewish low point of view.'
(3305)

In both (4a) and (4b) the events remain essentially the same as in (3a) and (3b). The difference lies in the construal of the situation, namely, in the profiled entities which the event is construed as related to. I will call this type of construal *a Dative construal of events*. In a Dative construal an event is being related to an external entity (marked by the Dative), which is profiled as an extra participant in the situation. Such a Dative construal can extend both affective (nonstative, dynamic) and non-affective (stative) situations. In a case of an affective situation, the Dative-marked participant is (prototypically) affected as a consequence of the situation described by the predicate and its core participants. A stative Dative construal is presented in sections 5.4–5.6, as Discourse Profile Constructions emerging from clusters Three, Four, and Five, respectively. The next section discusses Cluster One, defining the Extended transitive Discourse Profile Construction

5.2.2 Extension of an affective event

The Extended transitive Discourse Profile Construction is a usage pattern used to construe a relation between two participants that produces a high degree of effect on a third entity. The third entity in the current case can be analysed as a Recipient, an Affectee, or a Possessive Dative. It is a cover function unifying all Extensions of an affective situation, in which the core event can be construed as complete without any reference to a third participant. It is in the Extended transitive Discourse Profile Construction that a two-place event gains a relative, dependent status, anchored to a third, Dative-marked participant. For example, it is only in the current Discourse Profile Construction that the two-participant event of 'drinking water' is construed as related to a third entity, consequently affecting it:

(5) *holxim* ***lishtot li et ha-maym*** *me-ha-shorashim shel ha-cmaxim.*
walk to.drink to.me ACC the-water from-the-roots of the-plants.
'They're gonna drink my plants' water (= they are going to draw water from a nearby underground location, whereby my plants will loose their water supply).'
(882)

And in (6) the Dative-marked participant can be termed a Recipient according to traditional, participant role definitions: it is affected by a transitive motion event in which an Agent (metaphorically) moves a Theme from one point to another:

(6) ***atem*** *lo* ***notnim lo et ha-kelim*** *le-hitmodedut im ha-beaya*
you not give to.him ACC the-tools to-cope with the-problem
ha-zot.
the-this.
'You don't give him the right tools for coping with this problem.' (109)

However, it still involves partial affectedness and indirect relations: the Recipient does not go through a complete change of state. The only entity that goes through a change of state in the construed event is the transferred theme. The Recipient is affected only as a result of the Agent's action on a different object, to the extent that it becomes a possessor of the transferred theme. This type of relation is very basic, and thus lexically encoded by transfer verbs.

Recall that a cluster's unique exemplars are the ones that share very few features with other tokens in other clusters in the data. In fact, the most unique exemplars in a cluster include verbs with a lexical Dative complement (such as (6), for example, with the lexical-Dative verb *natan* 'give').

Thus, the prototype of the construction is a high transitivity, non-subjective clause, that construes a relation between a two participant event and a third participant. In the prototypical case, it is transfer relations that get construed. In the less prototypical cases, it can be any high transitivity event that can be construed as affecting a third participant. This construal is done through the Extended transitive Discourse Profile Construction. This description of Cluster One's prototypical exemplar is in line with Boneh and Bar-Asher Siegal's (2014) account of non-lexical Datives in Hebrew. The exact type of effect, and the particular type of participant role associated with the third participant (i.e., the Dative function) is dictated by both the predicate and the relations between the Direct Object referent and the Dative-marked participant. If a possession relation exists, for instance, it is the Malefactive–Possessive Dative role. If, however, the possession relation results from the event rather than from its circumstances, the participant role of the Dative-marked participant is the Benefactive.

5.3 Cluster Two

Cluster Two is the largest cluster in the data, grouping together 4,127 tokens of Dative clauses – 42.5% of the corpus. It is a coherent and homogeneous cluster, linked with transitive predicates from the types Primary B and Secondary C. Recall that Primary B verbs have concrete participants, but they can also be complemented by a clause (either instead of or in addition to the nominal complement), while Secondary C verbs introduce a subject that actually plays a role in bringing about the event or state referred to in the complement clause's verb. These predicates are situated in a three-argument structure, but unlike Cluster One, the Direct Object slot is occupied by a clause or a Preposition Phrase rather than a Noun Phrase. That is, two arguments are referential, while the third argument is a clausal complement of the predicate conveying the content in a telling or showing event (Primary B verbs), or a Prepositional complement of the predicate in case of Secondary C verbs. The verb paradigms (*binyan*) linked with Cluster Two are HIFIL, KAL and PIEL showing a weaker link. Cluster Two is also linked to the Irrealis mode, and to a Dative referent in the third person, which in participant role terms would be analyzed as one of the following: (1) an Addressee, the human endpoint of a message delivering action, its audience, or (2), a Human Endpoint, which is the partially/mentally affected endpoint of a two-participant event, marked as human (see Section 3.4).

These characteristics of Cluster Two can be exemplified by looking at its central exemplars – the tokens that share the largest number of features with other

tokens within the cluster – and its unique exemplars – the tokens that belong to Cluster Two and share the smallest number of features with tokens from other clusters. The central exemplars of Cluster Two construe an event in which the Dative-marked participant is the only endpoint of the action initiated by the A referent:

(7) *yesh xavery kneset she-**azru** **lanu** b-a-inyan ha-ze.*
there.are parliament members that-helped to.us in-the-issue the-this.
'Some parliament members helped us with that.'
(130)

Syntactically, the third argument is a Prepositional complement of the main predicate, usually indicating a secondary event the initiator of which is the Dative-marked participant. This construal is related to an intermediate degree of affectedness of the Dative-marked participant, as seen in Chapter 4, and the tokens of Cluster Two are located accordingly, mostly at the bottom-left quadrant of the Multiple Correspondence Analysis map.

The unique exemplar of Cluster Two, presented in (8) (sharing the smallest number of features with tokens belonging to other clusters in the dataset) construes a telling event in which the Dative-marked participant is mentally affected:

(8) *ani **macia** **lexa** ke-ohed **she-tisgor** oto le-reva shaa.*
I suggest to.you as-fan that-you.will.close it for-quarter hour.
'As a fan, I suggest that you close it for fifteen minutes'.
(30)

Note that the third argument in (8) is a finite clause indicating the content of a message delivered by the A referent to the Dative-marked participant. That is, in both (7) and (8) the Dative-marked participant is the only real affected-endpoint of the event, the clausal/Prepositional complement being non-referential (Thompson and Hopper, 2001). The construal of the telling event conveyed in (8) profiles the audience of the telling action (the Dative-marked participant) as the single endpoint of the action. This is in contrast with telling construals that belong to Cluster One, where the *type of message* is profiled and occupies a full nominal argument position, the O argument:

(9) *anaxnu poxadim **lehacia** **lahem** **hacaot**.*
we afraid to suggest to.them suggestions.
'We're afraid of giving them any suggestions.'
(349)

In the 'type of message' case, exemplified in (9), the Dative-marked participant (the audience) is not construed as the only affected entity in the event, but rather as indirectly affected by the metaphorical creation of a message. Thus, the telling act is construed differently when used within the usage patterns of Cluster One and Cluster Two: in the Discourse Profile Construction that emerges from the tokens of Cluster One a telling/showing event is construed as a creation of an abstract entity. This creation is done with relation to another entity, thus affecting it. In the Discourse Profile Construction emerging from the tokens of Cluster Two, however, a telling/showing event is construed as a two-pole action, with a teller/shower and an audience/perceiver, emphasizing not the creation of the message, but its content.

This is, then, the Discourse Profile Construction emerging from Cluster Two: a set of lexical, morpho-syntactic, semantic and pragmatic features that can be summarized as related to an intermediate level of transitivity and a rather high degree of subjectivity (see Chapter 4), conventionally paired with an interpretation of the Dative-marked participant as mildly/mentally affected by an action initiated by the A referent. This type of action is usually realized through either a telling/showing or helping/causing events as shown above, but not always, as can be seen in the following Cluster Two examples:

(10) a. *eyze metayel yirce lehagia le-shetax tiyuley ofnaym*
what traveller will want to arrive to-field trips.of bicycle
kshe-kodxim lo, *hu xoshesh mi-zihum svivati.*
when-drill to.him, he afraid of-pollution environmental.
'What traveller would like to have a bicycle trip in a drilling area? He's afraid it would be polluted.'
(1,457)

b. *ke-xol she-ze simeax oti, ze gam* ***hixiv li.***
as-all that-it made happy ACC.me, it also hurt to.me.
'As much as it made me happy, it hurt me too.'
(12,237)

c. *im misheu xayav kesef [...] ve-hu lo yaxol leshalem,*
if someone owes money [...] and-he not can to pay,
mevatrim lo *al ze.*
relinquish.PL to.him about it.
'If one owes money and they can't pay, their debt will be forgotten.'
(13,988)

What groups these exemplars together with the central and unique exemplars presented above is a small-scale, fine-grained family resemblance. That is, it is not a

case of metaphorical extension from *help* to *relinquish*, or to *drill*. Rather, these cases construe an event with an intermediate degree of transitivity in which the Dative-marked participant is the only endpoint, partially or mentally affected. These exemplars share more formal similarities with other exemplars belonging to Cluster Two than with exemplars from other clusters in the corpus: it is a combination of a transitive verb with a two-argument, Subject-first structure in an Irrealis clause, or a three-argument structure with a non-nominal occupying the third argument position (i.e., no individuated O), and an intransitive verb in the active voice, combined with a rather low degree of Agentivity of the A participant in the event.

Summing up the discussion regarding the second cluster in the data, we have seen that the organizing principle of the Affectedness Scale as presented in Chapter 4 comes about again in the grouping of different exemplars (which could have been analysed as different in traditional, participant role terms) according to their shared usage pattern. This cluster of tokens is an emergent Discourse Profile Construction which pairs this shared usage pattern with a particular construal of an event. In this construal of the events, the Dative-marked participant is the only affected entity; it is marked as undergoing an intermediate degree of affectedness, be it either physical but partial, or mental. Such a partial/mental effect is unique to human undergoers, and hence the Dative marking.

5.4 Cluster Three

Cluster Three is composed of 1,110 tokens of Dative clauses. These tokens are linked to the Secondary D, B, and A types of predicates (ordered according to strength of link), mostly in the NIFAL verb paradigm, and to the Primary A type in a weaker link. Recall that Secondary D verbs are intransitive verbs that take a complement clause in subject slot and another role which is some sort of stance taker, marked by the Dative. Secondary B verbs have one independent role (the subject of the Secondary-B verb) in addition to the roles of the verb in the complement clause describing the subject's attitude towards some event or state. Secondary A verbs have no independent semantic roles and modify the meaning of another verb (the main verb of a complement clause), sharing its roles and syntactic relations. And Primary A verbs have only concrete participants as their semantic roles. Cluster Three is linked to a two-argument, predicate-first, syntax, with the predicate either in the NIFAL paradigm, as mentioned above, or in the passive voice, thus appearing in the HUFAL and PUAL verb paradigms. The non-Dative argument most strongly linked to Cluster Three is of an Adjectival nature,

as in:

(11) *ze nire li* ***naxon***.
it seems to.me right.
'It seems right to me.'
(219)

It can, however, be a clause or a nominal as well. In nominal cases, it is low in Agentivity. The Dative referent linked to Cluster Three is in the first person, suggesting high Subjectivity, and it corresponds with the Experiencer or the Ethical Dative as has been analysed in Chapter 3. For example, consider the most central token in the cluster, presenting the Dative-marked participant as a cognizer, a role traditionally subsumed under the Experiencer Dative function:

(12) *me-olam lo yadati she-ani yaxol lemale et ze b-a-internet ad*
never not I.knew that-I can fill ACC this in-the-web until
she-ze ***noda*** ***li***.
that-it was.known to.me.
'I didn't know I can go on-line and fill these forms until it became known to me.'
(624)

While the Experiencer is the participant role most strongly linked with Cluster Three, other types of participant roles can be found in the tokens close to the center of the cluster as well. For instance, the Dative-marked participant in (13) can be analysed as a Recipient or a Possessive Dative:

(13) *hu yirce leharviax kesef, hu yirce sh-****yishaer*** ***lo***
he will.want to earn money, he will.want that-will.stay to.him
mashu b-a-yad.
something in-the-hand.
'He'll want to get something out of it, to earn some money, to have something in his hand.'
(1,078)

However, the Dative-marked participant in (13) is profiled as wishing to be in a state in which he has some money, therefore it is more of an Experiencer than a Possessor.

The most unique token in the cluster, on the other hand, includes an Addressee Dative-marked participant, which is the human endpoint of a message delivering action. In Cluster Three, however, such an Addressee, a Recipient or a Possessive

Dative are related either to intransitive predicates (as in (13)), or to the passive voice, as in:

(14) *hu kibel et ha-haca'a she-**huc'a** **lo**.*
he accepted ACC the-offer that-was.suggested to.him.
'He accepted the offer that was made to him.'
(10,730)

Considering this variety of examples, trying to characterize the tokens in Cluster Three as related to a single participant role is an impractical task. The question thus remains what brings these (and other) tokens together; i.e., what are the important similarities they feature, such that they belong to the same cluster. The tokens in Cluster Three are judged as similar not on the basis of a common participant role marked by the Dative, but rather on the basis of the construal conveyed by the speaker, realized as a shared usage pattern. That is, the Dative function and the overall construal of the clause are not necessarily dependent. The following sentences, for example, belong to Cluster Three as well, and include an intransitive verb, a predicate-first structure, and other low transitivity features as was discussed in Chapter 4.

(15) a. *ze yaxol likrot gam lahem.*
it can happen also to.them.
'It can happen to them as well.'
(7)

b. *ani agid laxem **ma** **kara** **li** ha-boker.*
I will.tell to.you what happened to.me this-morning.
'Let me tell you what happened to me this morning.'
(207)

c. *yesh mistanenim she-kvar hiclaxnu leharxik*
there.are infiltrators that-already we.succeeded to banish
*otam ve-**hem** **xazru** **lanu** paam nosefet.*
ACC.them and-they came.back to.us time another.
'There are some infiltrators that we have already banished, and they managed to come back.'
(517)

(16) a. *ze nire li kriti.*
it seems to.me critical.
'It seems critical to me.'
(428)

b. ***nire li*** *she-ze ha-davar ha-codek beyoter laasoto*
seems to.me that-it the-thing the-right most to.do.it
b-a-tvax ha-raxok.
in-the-range the-far.
'It seems to me to be the right thing to do in the long run.'
(542)

c. *mamash* ***lo ba li*** *she-hu yavo im ha-emda*
really not come to.me that-he will.come with the-viewpoint
ha-kodemet.
the-former.
'I would really hate it if he'd come here with his old viewpoint.'
(599)

This usage pattern is related to a construal that anchors a situation to a human referent, the Dative-marked participant. This is the *Extended Stative* Discourse Profile Construction. That is, summing up the discussion about the third cluster of tokens in the data, we can say that while a stative construal consists of an S argument and a predicate, the Extended Stative construal is composed of an S argument, a predicate, and an extra participant marked by the Dative (E, or extension to core, see Dixon (2005)).

5.5 Cluster Four

Cluster Four consists of 1,141 tokens, strongly linked to a two-argument syntactic structure with Adjectival predicates from the types Property and Value. As seen in Chapter 2, the Property type includes adjectives such as *barur*, 'clear,' *kashe*, 'hard,' and *nagish*, 'accessible.' The Value type includes adjectives such as *xashuv*, 'important,' *tov*, 'good,' and *nifla*, 'wonderful.' Another distinctive feature of Cluster Four's tokens is a predicate-first linear order, with the non-Dative argument realized as a non-finite clause. Cluster Four is linked to Dative referents in the first person, traditionally analysed as Experiencers, Judicantis, or Evaluative Reference Points in participant role terms (see Chapter 3). Consider the following sentences: the first two ((17a)–(17b)) are the closest to the center of gravity of the cluster (i.e., central exemplars), sharing the highest number of features with other exemplars in the cluster. The third and fourth examples ((17c)–(17d)) are the most unique exemplars of Cluster Four, sharing the smallest number of features with tokens from other clusters.

(17) a. *ma* ***laxuc*** ***lanu*** *laasot et ze axshav?*
what pressed to.us to do ACC it now?
'What's the rush to do it right now?'
(4,257)

b. ***barur li*** *laxalutin ha-racon legaven b-a-ir ha-zot.*
clear to.me totally the-desire to vary in-the-city the-this.
'I totally understand the desire to achieve a variation in this city.'
(1,971)

c. *im mankal [. . .] savur she-lo* ***rauy*** ***lo*** *lehagia*
if C.E.O [of. . .] thinks that-not appropriate to.him to arrive
l-a-diyun shel ha-kneset, ha-kneset af paam lo
to-the-discussion of the-Knesset, the-Knesset never not
neelevet.
gets insulted.
'If the C.E.O of . . . thinks the discussion is not important enough for him to be here, the Knesset never gets insulted.'
(14,658)

d. *lo* ***naxon lanu*** *leharim et ha-kfafa.*
not right to.us to raise ACC the-glove.
'We should not 'take up the glove'.'
(8,602)

Looking at these examples, and considering the Hierarchical Clustering on Principal Components output of the strength of links (see Appendix C), we can characterize Cluster Four as an Adjectival/Adverbial cluster. The tokens belonging to Cluster Four construe an evaluation of a state of affairs relative to, or from the point of view of, the Dative-marked participant. For example, the speaker in (17c) is not evaluating the arriving of the CEO as inappropriate; rather, this evaluation is anchored relative to a reference point – the Dative-marked participant. And in (17d), while the Dative-marked participant is in the first person plural (thus referring to the speaker), we can still see that there is no 'objective' evaluation of the situation as wrong, but rather, a subjective, relative reference to the situation, anchoring the evaluation to a reference point profiled by the Dative-marked participant. That is, in (17d) the conveyed meaning can be broken into two parts, much like the type of construal we have seen in the case of the Extended transitive Discourse Profile Construction emerging from Cluster One: an evaluation, and its anchoring to a reference point. And indeed, Cluster Four is the only cluster linked to what in participant role terms would be analysed as an Evaluative Reference Point.

In summary, the usage pattern of Cluster Four (of low affectedness, low transitivity, and high subjectivity, see Chapter 4) is conventionally paired with a construal of an evaluation of a state of affairs, anchored to the Dative-marked participant. This is, then, a Discourse Profile Construction.

5.6 Cluster Five

Cluster Five is the smallest cluster in the data, consisting of only 456 tokens. Clusters with a small number of tokens tend to be rather heterogeneous, and are less easily characterizable than other clusters that present a uniform behaviour. Cluster Five cannot be said to have a consistent type of predicate, for example; the predicate types strongly linked to Cluster Five according to the Hierarchical Clustering on Principal Components are Property and Value (Adjectival predicates), Secondary A, B, and D, Nominals, Adverbials, Prepositions (18a), an interrogative (18b), and Discourse Markers. The Dative functions associated with the tokens of Cluster Five vary as well, and it can refer to a Discourse Marker (18c), an Experiencer, or an Evaluative Reference Point. That is, there seems to be no coherent principle according to which tokens are categorized together into Cluster Five. Consider the cluster's central ((18a) and (18b)) and unique (18c) exemplars:

(18) a. *ani,* ***be-nigud lexa****, lo mitlahevet me-ha-nose.*
I, in-contrast to.you, not excited from-the-issue.
'Unlike you, I'm not so thrilled about this subject.'
(4,941)

b. ***lama lanu*** *lehaxbid?*
why to.us to make it heavy?
'We shouldn't make it harder than it is.'
(2,468)

c. ***toda raba lexa****, adoni ha-yoshev rosh.*
thank very much to.you, Mr. chairman.
'Thank you very much, Mr. chairman.'
(757)

The two central tokens of the cluster presented here do not seem to have a common function. The Dative clause in (18a) is not the main clause of the utterance, nor does it carry the utterance's main conveyed meaning. Rather, it is a parenthetical comment concerning the speech act and its interlocutors in this case. In fact, the Dative clause's function in the type of parenthetical comments involving a

pronominal Dative is to put the utterance in a relative, dependent state, anchoring it to an external entity – be it on of the interlocutors, or another relevant referent.

The second central exemplar of Cluster Five (18b) is different, and is part of the conveyed meaning rather than a discursive level parenthetical comment. It is a special use of an interrogative word in a construction that resembles the Discourse Profile Constructions we have seen so far, relating a question to a particular reference point. That is, a question that could have been impersonal and general as in the constructed example in (19), is subjectively construed as relevant only to the Dative-marked participant. We can call this Dative construal a *personalization construal*.

(19) *lama lehaxbid?*
why make it heavy?.
'Why make it difficult?' (or, 'One shouldn't make it difficult.')

The unique exemplar of Cluster Five (18c) is a gratitude utterance, aimed towards the Dative-marked participant. It resembles Dative clauses with other nominal predicates concerning manner and greetings (also belonging to Cluster Five) such as:

(20) a. *kol ha-kavod laxem.*
all the-respect to.you.
'Way to go!'
(2,432)
b. *shalom laxem.*
hello to.you.
'Hello!'
(2,704)

In some respects, these examples are not different from the Addressee participant role, or the Recipient. While in the Recipient cases it is a transfer event construal in which a transitive motion event is directed at the Dative-marked participant (section 3.5), and in the Addressee cases it is a telling/showing event (section 3.1), here it is a greeting action, lexicalized in a nominal rather than a verbal construction. Thus, there is no substantial difference between the nominal Dative clause in (18c) and its verb-headed equivalent in (21) that belongs to Cluster Two:

(21) *ani meod mode lexa.*
I very thank to.you.
'Thank you very much.'
(15,985)

The only difference is in construal. While in (21) the greeting action is construed as a transitive event consisting of an Initiator and an Endpoint, in (18c) the same action is construed as a stative situation with no explicit reference to the Initiator of the action, profiling its Endpoint as the only relevant human entity.

We can see, then, that there is no general Discourse Profile Construction that emerges from the tokens of Cluster Five, due to the cluster's heterogeneous nature. However, we can still detect a recurrent pattern of construals agreeing with the types of construals we have seen so far as related to Dative constructions.

Summing up the Hierarchical Clustering on Principal Components, in the present chapter we have shown that a set of Hebrew Dative clause types can be categorized into clusters in a bottom-up manner so that each cluster corresponds to a homogeneous type of construal. These construals have been defined as Discourse Profile Construction, a novel notion representing a conventional pairing of a multifactorial usage profile with a particular construal. The Hierarchical Clustering on Principal Components performed in the present chapter provided support for the findings from the Multiple Correspondence Analysis presented in the previous chapter. Each Discourse Profile Construction identified through the clustering analysis was shown to unify different Dative participant roles according to the level of affectedness of the Dative-marked participant.

6
Conclusions

The present work proposes a Usage-Based mapping of Hebrew Dative constructions. By means of corpus-based research, carried out within a usage-based perspective, I have considered not only syntactic or semantic features in isolation but taken into account multiple parameters from multiple sources of linguistic and extra-linguistic information. I have shown that the crucial key for the interpretation of a Dative clause is its location on an affectedness scale combining transitivity and subjectivity characteristics. That is, I claim that the interpretation of a Dative clause is a function of the complete construal, rather than traditionally defined participant roles such as an Experiencer, or a Benefactive. We have seen, for example, that while two clauses may share similar predicates, and a similar Dative-marked participant role, if their transitivity-subjectivity context is different they will convey a different degree of effect exerted upon the Dative-marked participant, profiling it in a different relation to the situation depicted by the clause, through a different construal of the world.

Tokens of Dative clauses are judged as similar or different not only on the basis of the lexical semantics of their predicate, but based on additional information as well. According to the findings reported here, the level of Agency of the A referent, the degree of Individuation of the O participant, the number and type of overt syntactic arguments in the clause, the type of verb paradigm (*binyan*), and the mode of the clause, all of these parameters in different constellations yield different groups of similar tokens, each group correlating with a particular

construal.

Using multivariate exploratory statistics we have been able to account for these similarities in an objective manner. This approach to similarity between tokens was supported by two complementary multivariate statistical tools. The first tool, Multiple Correspondence Analysis, provided insights regarding both the tokens in the data and the variables and categories they were coded for. That is, we were able to say that a set of tokens that share a particular category tend to share another set of categories as well. We have found consistency in the correspondences between categorical variables such that categories tend to correspond according to their shared level of transitivity- and subjectivity-related features. This consistent correspondence means that a set of categories is relevant for a particular, well defined, set of tokens. From the opposite point of view, tokens are judged as similar, located adjacent to each other on the Multiple Correspondence Analysis map, if they share enough unique categories.

The correspondences we have found between categories, and the sets of tokens they group together, were shown to correlate with sets of similar Dative functions (i.e., participant roles), converged into a single Dative construal. That is, while the Dative functions themselves were not part of the Multiple Correspondence Analysis calculation, the resulted map presented a division of the tokens in the corpus in a way almost identical to a traditional, participant role oriented division. However, the substantial differences were shown to exist between the aforementioned sets of converged Dative functions, rather than between locally defined participant roles such as an Experiencer or a Recipient.

Since the results of the Multiple Correspondence Analysis suggested a division of the corpus into several sets of tokens correlating with sets of converged Dative functions, I used a clustering process in order to quantify similarities, in an attempt to show that a bottom-up clustering of the tokens in the data will result with the same, visually and subjectively detected groups of tokens. The chosen tool for this job was Hierarchical Clustering on Principal Components. The result of the Hierarchical Clustering on Principal Components aided us in reaching a better (and more accurate) description of clusters of tokens that share a set of similar Dative functions. These clusters were described as a unique type of construction, and a novel theoretical concept, the Discourse Profile Construction. A Discourse Profile Construction was defined as a usage profile consisting of multifactorial, multilevel information, conventionally paired with a single construal of a situation. Thus, rather than accounting for different Dative functions, showing their subtle variations, Discourse Profile Constructions allow one to account for converged Dative functions corresponding to a single construal of the world. The interpretation of the Hierarchical Clustering on Principal Components yielded

four such Discourse Profile Constructions:

1. The Extended Transitive Discourse Profile Construction,
2. The Human Endpoint Discourse Profile Construction,
3. The Extended Intransitive Discourse Profile Construction,
4. The Evaluative Reference Point Discourse Profile Construction.

Each Discourse Profile Construction was defined based on the categories most strongly linked to it, and based on its central and unique exemplars. We have seen that the similarity between different tokens belonging to the same Discourse Profile Construction can be explained through a small-scale, exemplar-based, family resemblance, without seeking explanations in metaphorical extensions.

The Hierarchical Clustering on Principal Components, then, provided an explanation for the principles that categorize tokens together into a single cluster. That is, using a clustering process, and based on the findings from the Multiple Correspondence Analysis, I have succeeded in quantifying similarities and dissimilarities between Dative clauses. The four Discourse Profile Constructions have emerged from the data, representing different types of construals. Although these are different construals, they all share the characteristic of personalizing a situation, anchoring it to (what could be perceived as) an external participant. Indeed, this is the essence of the Dative construal.

Bibliography

Abbot-Smith, Kirsten, and Michael Tomasello. 2006. Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review* 23:275–290.

Abdi, Hervé, and Dominique Valentin. 2007. Multiple correspondence analysis. *Encyclopedia of measurement and statistics* 651–657.

Amritavalli, Raghavachari. 2004. Experiencer Datives in Kannada. In *Non-nominative Subjects*, ed. Peri Bhaskararao and Karumuri Venkata Subbarao, 1–24. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Ariel, Mira. 2008. *Pragmatics and Grammar*. Cambridge: Cambridge University Press.

Ariel, Mira, Elitzur Dattner, John W Du Bois, and Tal Linzen. In progress. Pronominal datives: The royal road to argument status.

Baayen, Harald R. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.

Baker, Collin F., Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet Project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 1*, ACL '98, 86–90. Stroudsburg, PA, USA: Association for Computational Linguistics. URL `http://dx.doi.org/10.3115/980845.980860`.

Beavers, John. 2011. On affectedness. *Natural Language & Linguistic Theory* 1–36.

Berman, Ruth A. 1982. Dative Marking of the Affectee Role. *Hebrew Annual Review* 6:35–59.

Blume, Kerstin. 1998. A contrastive analysis of interaction verbs with dative complements. *Linguistics* 36:253–280.

Boneh, Norah, and Elitzur Bar-Asher Siegal. 2014. Non obligatory Datives in Hebrew in their context In Hebrew.

Borer, Hagit, and Yosef Grodzinsky. 1986. Syntactic Cliticization and Lexical Cliticization: The Case of Hebrew Dative Clitics in The Syntax of Pronominal Clitics. In *Syntax and Semantics 19*, ed. Hagit Borer, 175–217. New York: Academic press.

Bybee, Joan L. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82:529–551.

Bybee, Joan L. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, Joan L., and David Eddington. 2006. A Usage-Based Approach to Spanish Verbs of 'Becoming'. *Language* 82:323–355.

Croft, William. 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago and London: The Universitiy of Chicago Press.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford; New York: Oxford University Press.

Croft, William. 2012. *Verbs: Aspect and Causal Structure*. Oxford University Press.

Croft, William, Jóhanna Barðdal, Willem Hollmann, Maike Nielsen, Violeta Sotirova, and Chiaki Taoka. 2001. Discriminating verb meanings: The case of transfer verbs. In *handout, LAGB Autumn Meeting, Reading*.

Cuervo, María Cristina. 2003. Datives at Large. Doctoral Dissertation, Massachusetts Institute of Technology.

Dabrowska, Ewa. 1994. Some English Equivalents of Polish Dative Constructions. *Papers and Studies in Contrastive Linguistics* 29:105–121.

Dabrowska, Ewa. 1997. *Cognitive Semantics and the Polish Dative*. Berlin/New York: Mouton de Gruyter.

Dattner, Elitzur. 2008. Lower Transitivity Constructions in Hebrew: The Case of Motion Verbs. Master's thesis, Tel Aviv University, Tel Aviv.

Diessel, Holger. 2004. *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.

Dixon, Robert M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Dixon, Robert M.W. 2005. *A semantic approach to English grammar*. Oxford: Oxford University Press, 2nd edition.

Dixon, Robert M.W. 2006. Complement clauses and complementation strategies in typological perspective. In *Complementation: a cross-linguistic typology*, ed. Robert M.W. Dixon and Alexandra Y. Aikhenvald, 1–48. Oxford: Oxford University Press.

Dowty, David. 1991. Thematic Proto-Roles and Argument Selection. *Language* 67:547–619.

Elman, Jeffrey L. 2009. On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive science* 33:547–582.

Even-Shoshan, Avraham. 2003. *Milon Even-Shoshan: Updated for the twenty first century: The comprehensive Hebrew lexicon*. [C.M]: Hamilon Haxadash.

Fillmore, Charles J., Paul Kay, and Mary Catherine O'Connor. 1988. Regularity and Idiomaticity in Grammatical Constructions: The Case of 'Let Alone'. *Language* 64:501–538.

Francez, Itamar. 2006. Possessors, goals, and the classification of ditransitive predicates: Evidence from Hebrew. *Empirical issues in Syntax and Semantics* 6:137–154.

Friendly, Michael. 1992. Graphical methods for categorical data. In *SAS User Group International Conference Proceedings*, volume 17, 190–200.

Gafter, Roey J. 2014. The Distribution of the Hebrew Possessive Dative Construction. guided by unaccusativity or prominence? *Linguistic Inquiry* 45:482–500.

Gahl, Susanne, and Alan C. L Yu. 2006. Introduction to the special issue on exemplar-based models in linguistics. *The Linguistic Review* 23:213–216. URL `http://www.reference-global.com/doi/abs/10.1515/TLR.2006.007`.

Glynn, Dylan. 2012. Correspondence analysis. An exploratory technique for identifying usage patterns. In *Corpus Methods in Cognitive Semantics*, ed. Dylan Glynn and Justyna Robinson, 133–179. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Glynn, Dylan, and Kerstin Fischer, ed. 2010. *Quantitative methods in cognitive semantics: Corpus-driven approaches*. Cognitive Linguistics Research. Berlin/New York: Walter de Gruyter.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University Of Chicago Press.

Goldberg, Adele E. 2003. Constructions: a New Theoretical Approach to Language. *Trends in Cognitive Sciences* 7:219–224.

Goldberg, Adele E. 2005a. Argument Realization: the role of constructions, lexical semantics and discourse factors. In *Construction Grammars. Cognitive grounding and theoretical extensions*, ed. Jan-Ola Östman and Mirjam Fried, 17–43. John Benjamins Publishing Company.

Goldberg, Adele E. 2005b. Constructions, Lexical Semantics, and the Correspondence Principle. In *The Syntax of Aspect: Deriving Thematic and Aspectual Interpretation*, ed. Nomi Erteschik-Shir and Tova Rapoport. Oxford: Oxford University Press.

Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, Adele E., and Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language* 80:532–568.

Greenacre, Michael, and Jörg Blasius. 2006. *Multiple Correspondence Analysis and Related Methods*. Boca Raton: Chapman & Hall/CRC.

Gries, Stefan Th. 2009. *Quantitative Corpus Linguistics With R: A Practical Introduction*. New York, NY: Routledge.

Gries, Stefan Th. 2013. *Statistics for linguistics with R: a practical introduction*. Berlin/Boston: Walter de Gruyter, 2nd revised edition.

Grimshaw, Jane, and Armin Mester. 1988. Light Verbs and Theta–Marking. *Linguistic Inquiry* 19:205–232.

Haspelmath, Martin. 2001. Non-Canonical Marking of Core Arguments in European Languages. In *Non-Canonical Marking of Subjects and Objects*, ed. Aikhenvald, Alexandra Y. and Dixon, Robert Malcolm Ward and Onishi, Masayuki, 53–83. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Haspelmath, Martin. 2003. The Geometry of Grammatical Meaning: Semantic Maps and Cross-linguistic Comparison. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, ed. Michael Tomasello, volume 2, 211–242. Mahwah, N.J.; London: Lawrence Erlbaum Associates.

Haspelmath, Martin. 2011. Ditransitive Constructions: The Verb 'Give'. In *The World Atlas of Language Structures Online*, ed. Matthew S. Dryer and Martin Haspelmath. Munich: Max Planck Digital Library. URL `http://wals.info/chapter/105`.

Hole, Daniel, André Meinunger, and Werner Abraham. 2006. *Datives and Other Cases: Between argument structure and event structure*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Hopper, Paul J., and Sandra A. Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56:251–299.

Husson, François, Sébastien Lê, and Jérôme Pagès. 2011. *Exploratory Multivariate Analysis by Example Using R*. Boca Raton: Chapman & Hall/CRC.

Husson, Francois, Julie Josse, Sebastien Le, and Jeremy Mazet. 2013. *FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R*. R package version 1.25.

Kay, Paul. 1997. *Words and the Grammar of Context*. Stanford, Calif.: CSLI Publications.

Kay, Paul, and Charles J. Fillmore. 1999. Grammatical Constructions and Linguistic Generalizations: The What's X Doing Y? Construction. *Language* 75:1–33.

Kemmer, Suzanne, and Michael Barlow. 2000. Introduction: A usage–based Conception of Language. In *Usage–Based Models of Language*, ed. Suzanne Kemmer and Michael Barlow. Stanford, Calif.: CSLI Publications.

Kishimoto, Hideki. 2010. The semantic basis of Dative case marking in Japanese. *Kobe Papers in Linguistics* 7:19–39.

Landau, Idan. 2010. *The Locative Syntax of Experiencers*. Cambridge, MA: MIT Press.

Langacker, Ronald W. 1990. Subjectification. *Cognitive Linguistics* 1:5–38.

Langacker, Ronald W. 2009. *Investigations in cognitive grammar*. Berlin/New York: Mouton de Gruyter.

Langendonck, Willy Van. 1998. The dative in Latin and the indirect object in Dutch. In *The Dative*, ed. William Van Belle and Willy Van Langendonck, volume 2, 211–260. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Lehmann, Christian. 2006. Participant roles, thematic roles and syntactic relations. In *Voice and grammatical relations: in honor of Masayoshi Shibatani*, ed. Tasaku Tsunoda and Taro Kageyama, 153–174. John Benjamins Publishing Company.

Levin, Beth. 1999. Objecthood: An event structure perspective. *Proceedings of CLS* 35:223–247.

Levin, Beth. 2008. Dative verbs: A crosslinguistic perspective. *Lingvisticæ Investigationes* 31:285–312.

Levin, Beth, and Malka Rappaport Hovav. 2005. *Argument Realization*. Cambridge: Cambridge University Press.

Levshina, Natalia. 2012. Comparing constructicons: A usage-based analysis of the causative construction with doen in Netherlandic and Belgian Dutch. *Constructions and Frames* 4:76–101.

Levshina, Natalia, Dirk Geeraerts, and Dirk Speelman. 2013. Mapping constructional spaces: A contrastive analysis of English and Dutch analytic causatives. *Linguistics* 51:825–854.

Linzen, Tal. 2009. Hebrew Possessive Datives: The effects of affectedess. Master's thesis, Tel Aviv University.

Lyngfelt, Benjamin, and Torgrim Solstad, ed. 2006. *Demoting the Agent: Passive, Middle and Other Voice Phenomena*. John Benjamins Publishing Company.

Malchukov, Andrej, Martin Haspelmath, and Bernard Comrie, ed. 2010. *Studies in Ditransitive Constructions. A Comparative Handbook*. Berlin: Mouton de Gruyter.

Nenadic, Oleg, and Michael Greenacre. 2007. Correspondence Analysis in R, with two- and three-dimensional graphics: The ca package. *Journal of Statistical Software* 20:1–13.

Newman, John. 1998. Recipients and 'give' constructions. In *The Dative*, ed. William Van Belle and Willy Van Langendonck, volume 2, 1–28. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Næss, Åshild. 2007. *Prototypical Transitivity*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rappaport Hovav, Malka, and Beth Levin. 2008. The English Dative Alternation: The Case for Verb Sensitivity . *Journal of Linguistics* 44:129–167.

Reinhart, Tanya. 2003. The theta system–an overview. *Theoretical linguistics* 28:229–290.

Rencher, Alvin. 2012. *Methods of multivariate analysis*. Hoboken, New Jersey: Wiley.

Shibatani, Masayoshi. 2001. Non-Canonical Constructions in Japanese. In *Non-Canonical Marking of Subjects and Objects*, ed. Aikhenvald, Alexandra Y, and Dixon, R.M.W, and Onishi, Masayuki. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Smith, Tomoko Yamashita. 2010. Cross-linguistic categorization of benefactives by event structure. In *Benefactives and Malefactives: Typological Perspectives and Case Studies*, ed. F. Zúñiga and S. Kittilä, 71–96. Amsterdam/Philadelphia: John Benjamins.

Sridhar, Shikaripur N. 1979. Dative Subjects and the Notion of Subject. *Lingua* 49:99–125.

Stefanowitsch, Anatol, and Stefan Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8:209–243.

Thompson, Sandra A., and Paul J. Hopper. 2001. Transitivity, Clause Structure, and Argument Structure: Evidence from Conversation. In *Frequency and the Emergence of Linguistic Structure*, ed. Joan L. Bybee and Paul J. Hopper, 27–60. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of Linguistics* 385–396.

Van Belle, William, and Willy Van Langendonck, ed. 1996. *The Dative*, volume 1: Descriptive Studies. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Van Langendonck, Willy, and William Van Belle, ed. 1998. *The Dative*, volume 2: Theoretical and Contrastive Studies. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Verhagen, Arie. 2005. *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition*. Oxford: Oxford University Press.

Šaríc, Ljiljana. 2002. On the Semantics of the 'Dative of Possession' in the Slavic Languages: An Analysis on the Basis of Russian, Polish, Croatian/Serbian and Slovenian Examples. *Glossos* 3.

von Waldenfels, Ruprecht. 2012. *The Grammaticalization of 'Give'+ Infinitive: A Comparative Study of Russian, Polish, and Czech*. Berlin/Boston: Walter de Gruyter.

Wedel, Andrew B. 2006. Exemplar models, evolution and language change. *The linguistic review* 23:247—274.

Zúñiga, Fernando, and Seppo Kittilä. 2010. Introduction. In *Benefactives and Malefactives: Typological perspectives and case studies*, ed. F. Zúñiga and S. Kittilä, volume 92. Amsterdam/Philadelphia: John Benjamins Publishing Company.

A

Frequency of predicates

Table A.1: Types and frequency of predicates in the corpus

Type	Predicate	Frequency
Primary A	*natan*, ‘give’	1,204
Primary A	*heevir*, ‘transfer’	94
Primary A	*hevi*, ‘bring’	80
Primary A	*lakax*, ‘take’	78
Primary A	*shilem*, ‘pay’	74
Primary A	*karah*, ‘happen’	67
Primary A	*hicig*, ‘present’	58
Primary A	*nitan*, ‘be given’	57
Primary A	*salax*, ‘forgive’	53
Primary A	*ala*, ‘rise, cost’	47
Primary A	*hexzir*, ‘return’	45
Primary A	*shalax*, ‘send’	44
Primary A	*sipek*, ‘supply’	32
Primary A	*hosif*, ‘add’	28
Primary A	*yaca*, ‘got to do’	26
Primary A	*hexin*, ‘prepare’	25
Primary A	*higish*, ‘serve’	25

Continued on next page

Table A.1 – *Continued from previous page*

Type	**Predicate**	**Frequency**
Primary A	*patax*, ‘open’	25
Primary A	*teer*, ‘describe’	25
Primary A	*nishar*, ‘be left’	24
Primary A	*hishir*, ‘leave’	21
Primary A	*hikshiv*, ‘listen’	20
Primary A	*mukar*, ‘known’	18
Primary A	*horid*, ‘reduce’	17
Primary A	*hikri*, ‘read aloud’	15
Primary A	*yicer*, ‘produce’	15
Primary A	*hoci*, ‘take out’	14
Primary A	*taram*, ‘donate’	14
Primary A	*sam*, ‘put’	13
Primary A	*bana*, ‘build’	12
Primary A	*haras*, ‘destroy’	12
Primary A	*maxar*, ‘sell’	12
Primary A	*hucag*, ‘be presented’	11
Primary A	*patar*, ‘solve’	11
Primary A	*hikdish*, ‘dedicate’	10
Primary A	*kana*, ‘buy’	9
Primary A	*rasham*, ‘write, note’	9
Primary A	*xayav*, ‘owe’	9
Primary A	*xipes*, ‘search’	9
Primary A	*nogea*, ‘concern’	8
Primary A	*xilek*, ‘divide, hand out’	8
Primary A	*yixes*, ‘relate’	8
Primary A	*himxish*, ‘illustrate’	7
Primary A	*huxzar*, ‘be returned’	7
Primary A	*mal’u*, ‘to reach a certain age’	7
Primary A	*xasax*, ‘save’	7
Primary A	*dafak*, ‘knock, bang’	6
Primary A	*histader*, ‘work out’	6
Primary A	*hoil*, ‘be useful’	6
Primary A	*hugash*, ‘be served’	6
Primary A	*niftax*, ‘be opened’	6
Primary A	*nitek*, ‘disconnect’	6
Primary A	*peret*, ‘describe, detail, elaborate’	6
Primary A	*shamar*, ‘save’	6

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
Primary A	*xatam*, ‘sign’	6
Primary A	*xidesh*, ‘Re-establish, renewed’	6
Primary A	*amad*, ‘stand’	5
Primary A	*arax*, ‘organize’	5
Primary A	*hicmid*, ‘attach’	5
Primary A	*hikna*, ‘instill’	5
Primary A	*hirbic*, ‘hit’	5
Primary A	*hishlim*, ‘complete’	5
Primary A	*huanak*, ‘be granted’	5
Primary A	*huavar*, ‘be transfered’	5
Primary A	*kafac*, ‘jump’	5
Primary A	*sagar*, ‘close’	5
Primary A	*shalal*, ‘rule out’	5
Primary A	*shimesh*, ‘be used’	5
Primary A	*shina*, ‘change, matter’	5
Primary A	*shulam*, ‘be paid’	5
Primary A	*yashav*, ‘sit’	5
Primary A	*acar*, ‘stop’	4
Primary A	*badak*, ‘check’	4
Primary A	*barax*, ‘escape’	4
Primary A	*bitel*, ‘cancel’	4
Primary A	*ganav*, ‘steal’	4
Primary A	*halax*, ‘go’	4
Primary A	*heela*, ‘raise’	4
Primary A	*hexlif*, ‘exchange’	4
Primary A	*hexmi*, ‘flatter’	4
Primary A	*hidbik*, ‘glue’	4
Primary A	*hikciv*, ‘allocate’	4
Primary A	*hixpil*, ‘multiply’	4
Primary A	*hukca*, ‘be allocated’	4
Primary A	*nafal*, ‘fall’	4
Primary A	*nihel*, ‘manage’	4
Primary A	*nishkaf*, ‘be foreseen’	4
Primary A	*nixnas*, ‘enter’	4
Primary A	*xazar*, ‘return’	4
Primary A	*yarad*, ‘descend’	4
Primary A	*yuxas*, ‘be related’	4

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
Primary A	*zarak*, 'throw'	4
Primary A	*arev*, 'guarantee'	3
Primary A	*citet*, 'quote'	3
Primary A	*darush*, 'be needed'	3
Primary A	*gamar*, 'finish'	3
Primary A	*hafax*, 'turn over'	3
Primary A	*higdir*, 'define'	3
Primary A	*hikca*, 'allocate'	3
Primary A	*himci*, 'invent'	3
Primary A	*hinpik*, 'issue'	3
Primary A	*hishmia*, 'make a sound'	3
Primary A	*hitarev*, 'interrupt'	3
Primary A	*hitpocec*, 'explode'	3
Primary A	*hitvasef*, 'be added'	3
Primary A	*hixnis*, 'insert'	3
Primary A	*hizik*, 'harm'	3
Primary A	*hofia*, 'appear'	3
Primary A	*hugdar*, 'be defined'	3
Primary A	*hukcav*, 'be allocated'	3
Primary A	*hukna*, 'be instilled'	3
Primary A	*husaf*, 'be added'	3
Primary A	*kara'*, 'tear'	3
Primary A	*laag*, 'mock'	3
Primary A	*mile*, 'fill'	3
Primary A	*nocar*, 'be created'	3
Primary A	*pina*, 'evacuate'	3
Primary A	*pocec*, 'destroy'	3
Primary A	*tirgem*, 'translate'	3
Primary A	*xulak*, 'be divided, handed out'	3
Primary A	*yaca*, 'exit'	3
Primary A	*yiker*, 'raise the price'	3
Primary A	*af*, 'fly'	2
Primary A	*avad*, 'got lost'	2
Primary A	*avar*, 'pass'	2
Primary A	*bilbel*, 'confuse'	2
Primary A	*ciyer*, 'paint'	2
Primary A	*daxa*, 'postpone'	2

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
Primary A	*gilgel*, 'roll'	2
Primary A	*heemid*, 'position'	2
Primary A	*heenik*, 'grant'	2
Primary A	*heezin*, 'listen attentively'	2
Primary A	*hefic*, 'distribute'	2
Primary A	*hegiv*, 'respond'	2
Primary A	*heir*, 'illuminate'	2
Primary A	*heniax*, 'leave alone, set down'	2
Primary A	*hidgim*, 'demonstrate'	2
Primary A	*hikir*, 'acknowledge'	2
Primary A	*hikrin*, 'project'	2
Primary A	*hilva*, 'lend'	2
Primary A	*hishtalet*, 'gain control'	2
Primary A	*histakel*, 'look'	2
Primary A	*hukdash*, 'be dedicated'	2
Primary A	*hunpak*, 'be issued'	2
Primary A	*ikev*, 'delay'	2
Primary A	*kadax*, 'drill'	2
Primary A	*kicec*, 'cut off'	2
Primary A	*na'asa*, 'be made'	2
Primary A	*neelam*, 'disappear'	2
Primary A	*neetar*, 'accede'	2
Primary A	*nolad*, 'born'	2
Primary A	*paga*, 'harm'	2
Primary A	*palash*, 'invade'	2
Primary A	*safar*, 'count'	2
Primary A	*saraf*, 'burn'	2
Primary A	*shar*, 'sing'	2
Primary A	*shata*, 'drink'	2
Primary A	*shixrer*, 'release'	2
Primary A	*sider*, 'organize, fix'	2
Primary A	*supak*, 'be supplied'	2
Primary A	*tipel*, 'take care of'	2
Primary A	*xared*, 'scared'	2
Primary A	*xata*, 'sin'	2
Primary A	*zar*, 'strange'	2
Primary A	*asaf*, 'gather'	1

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
Primary A	*baxa*, ‘cry’	1
Primary A	*baxan*, ‘examine’	1
Primary A	*baxar*, ‘choose’	1
Primary A	*benigud*, ‘as opposed to’	1
Primary A	*ceref*, ‘add’	1
Primary A	*cilem*, ‘photograph’	1
Primary A	*cimcem*, ‘reduce’	1
Primary A	*dalaf*, ‘leak’	1
Primary A	*gadal*, ‘grow’	1
Primary A	*gamal*, ‘recompense’	1
Primary A	*harag*, ‘kill’	1
Primary A	*hecif*, ‘flood’	1
Primary A	*heelim*, ‘conceal’	1
Primary A	*heerix*, ‘lengthen, extend ’	1
Primary A	*hefik*, ‘produce’	1
Primary A	*hekim*, ‘establish’	1
Primary A	*herim*, ‘lift’	1
Primary A	*hexdir*, ‘insert’	1
Primary A	*hexrim*, ‘confiscate’	1
Primary A	*hexshiv*, ‘ascribe inportance’	1
Primary A	*heziz*, ‘move something’	1
Primary A	*hicbia*, ‘vote’	1
Primary A	*hicdia*, ‘salute’	1
Primary A	*hiciv*, ‘place’	1
Primary A	*hictaber*, ‘accumulate’	1
Primary A	*hictanea*, ‘be modest’	1
Primary A	*hifna*, ‘turn’	1
Primary A	*hifxit*, ‘subtract’	1
Primary A	*higdil*, ‘cause to grow’	1
Primary A	*hikir*, ‘introduce’	1
Primary A	*hikpi*, ‘freeze’	1
Primary A	*hilbin*, ‘whiten, bleach’	1
Primary A	*hipil*, ‘drop’	1
Primary A	*hir’il*, ‘poison’	1
Primary A	*hishpia*, ‘influence’	1
Primary A	*hishtolel*, ‘be unruly’	1
Primary A	*hishxit*, ‘vandalize’	1

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
Primary A	*hisig*, ‘accomplish’	1
Primary A	*hisvev*, ‘turn around’	1
Primary A	*hithapex*, ‘turn over’	1
Primary A	*hitna*, ‘condition’	1
Primary A	*hitnaged*, ‘resist’	1
Primary A	*hitnaker*, ‘ignore, alienate’	1
Primary A	*hitparec*, ‘burst in’	1
Primary A	*hitva*, ‘outline’	1
Primary A	*hitxaber*, ‘become attached’	1
Primary A	*hitxalef*, ‘be changed’	1
Primary A	*hitxaxesh*, ‘deny’	1
Primary A	*hivrik*, ‘shine’	1
Primary A	*hix’iv*, ‘hurt’	1
Primary A	*hizrik*, ‘inject’	1
Primary A	*hizrim*, ‘cause flow’	1
Primary A	*hokir*, ‘cherish’	1
Primary A	*hudbak*, ‘be glued’	1
Primary A	*hurad*, ‘be reduced’	1
Primary A	*hutal*, ‘be imposed’	1
Primary A	*huxan*, ‘be prepared’	1
Primary A	*ifyen*, ‘characterize’	1
Primary A	*ir’er*, ‘undercut, destabilize’	1
Primary A	*irgen*, ‘organize’	1
Primary A	*kacav*, ‘allocate’	1
Primary A	*kam*, ‘get up’	1
Primary A	*karac*, ‘wink’	1
Primary A	*karat*, ‘cut down’	1
Primary A	*karav*, ‘approach’	1
Primary A	*kasam*, ‘enchant’	1
Primary A	*kata*, ‘interrupt’	1
Primary A	*katua*, ‘be disrupted’	1
Primary A	*kicer*, ‘shorten’	1
Primary A	*kilkel*, ‘damage’	1
Primary A	*kisa*, ‘cover’	1
Primary A	*kiseax*, ‘hit violently’	1
Primary A	*kizez*, ‘deduct’	1
Primary A	*laxash*, ‘whisper’	1

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
Primary A	*laxuc*, ‘pressed, stressed’	1
Primary A	*mana*, ‘count’	1
Primary A	*maxak*, ‘erase’	1
Primary A	*maxal*, ‘forgive’	1
Primary A	*mecoraf*, ‘attached’	1
Primary A	*misa*, ‘tax’	1
Primary A	*naga*, ‘touch’	1
Primary A	*neecar*, ‘be delayed’	1
Primary A	*ne’erax*, ‘be organized’	1
Primary A	*neheras*, ‘be destroyed’	1
Primary A	*nexsam*, ‘be blocked’	1
Primary A	*nicbar*, ‘be collected’	1
Primary A	*nidbak*, ‘be glued’	1
Primary A	*nidlak*, ‘lit, to be lighted’	1
Primary A	*nifga*, ‘be injured’	1
Primary A	*nifnef*, ‘wave’	1
Primary A	*niftar*, ‘be solved’	1
Primary A	*nigmar*, ‘over’	1
Primary A	*nigzar*, ‘be sentenced’	1
Primary A	*nika*, ‘deduct’	1
Primary A	*nikev*, ‘puncture’	1
Primary A	*nilkax*, ‘be taken’	1
Primary A	*nisgar*, ‘be closed’	1
Primary A	*nitpas*, ‘be caught’	1
Primary A	*nivna*, ‘be built’	1
Primary A	*nixtav*, ‘be writen’	1
Primary A	*nixtav*, ‘written’	1
Primary A	*nizkak*, ‘need’	1
Primary A	*parac*, ‘break in’	1
Primary A	*pasak*, ‘rule’	1
Primary A	*perek*, ‘dissemble’	1
Primary A	*pizer*, ‘scatter’	1
Primary A	*prat*, ‘excluding’	1
Primary A	*rikez*, ‘assemble’	1
Primary A	*roked*, ‘dance’	1
Primary A	*satam*, ‘seal’	1
Primary A	*saxar*, ‘hire’	1

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
Primary A	*shafax*, 'spill'	1
Primary A	*shamat*, 'drop'	1
Primary A	*shatal*, 'plant'	1
Primary A	*shavar*, 'break'	1
Primary A	*shibesh*, 'disturb'	1
Primary A	*shidreg*, 'upgrade'	1
Primary A	*shikef*, 'reflect'	1
Primary A	*shiker*, 'lie'	1
Primary A	*shiklel*, 'weigh'	1
Primary A	*shiyex*, 'attribute'	1
Primary A	*shudax*, 'be paired'	1
Primary A	*siber*, 'give a simple explanation'	1
Primary A	*sibsed*, 'subsidize'	1
Primary A	*sikem*, 'summarize'	1
Primary A	*tafas*, 'catch'	1
Primary A	*talash*, 'detach'	1
Primary A	*taman*, 'hide'	1
Primary A	*tava*, 'drown'	1
Primary A	*tiftef*, 'drip'	1
Primary A	*tikcev*, 'allocate'	1
Primary A	*tiken*, 'fix, amend'	1
Primary A	*utar*, 'be located'	1
Primary A	*xakar*, 'investigate'	1
Primary A	*xarat*, 'engrave'	1
Primary A	*xiber*, 'join'	1
Primary A	*yied*, 'designate'	1
Primary A	*yixed*, 'dedicate'	1
Primary B	*amar*, 'say'	1,435
Primary B	*higid*, 'tell'	532
Primary B	*hoda*, 'thank'	277
Primary B	*siper*, 'tell'	141
Primary B	*kara*, 'name'	131
Primary B	*hicia*, 'suggest'	115
Primary B	*hisbir*, 'explain'	113
Primary B	*hera*, 'show'	101
Primary B	*yadua*, 'known'	94
Primary B	*ana*, 'answer'	86

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
Primary B	*hodia*, 'inform'	84
Primary B	*hivtiax*, 'promise'	64
Primary B	*hizkir*, 'remind'	51
Primary B	*kara*, 'call'	46
Primary B	*katav*, 'write'	39
Primary B	*masar*, 'deliver'	31
Primary B	*nimsar*, 'be delivered'	30
Primary B	*diveax*, 'report'	24
Primary B	*maca*, 'find'	23
Primary B	*hivhir*, 'clarify'	19
Primary B	*kava*, 'determine'	19
Primary B	*da*, 'know'	18
Primary B	*gila*, 'discover'	16
Primary B	*histaber*, 'become clear'	16
Primary B	*hitbarer*, 'become clear'	16
Primary B	*noda*, 'become known'	16
Primary B	*heshiv*, 'return, answer, reply'	13
Primary B	*kara*, 'read'	13
Primary B	*hoxiax*, 'prove'	12
Primary B	*nikba*, 'be decided'	12
Primary B	*zaxur*, 'be remembered'	12
Primary B	*nimca*, 'be found'	9
Primary B	*biser*, 'inform'	8
Primary B	*heir*, 'remark'	8
Primary B	*sider*, 'organize, fix'	7
Primary B	*caak*, 'shout'	5
Primary B	*muvan*, 'understood'	5
Primary B	*rashum*, 'written'	5
Primary B	*katuv*, 'written'	4
Primary B	*naana*, 'be accepted'	4
Primary B	*shama*, 'hear'	4
Primary B	*ciyen*, 'note'	3
Primary B	*raxash*, 'feel (respect)'	3
Primary B	*duvax*, 'be reported'	2
Primary B	*hitgala*, 'be discovered'	2
Primary B	*hitif*, 'preach'	2
Primary B	*katuv*, 'writen'	2

Continued on next page

Table A.1 – *Continued from previous page*

Type	**Predicate**	**Frequency**
Primary B	*ba-beteana*, 'complain'	1
Primary B	*her'a*, 'show'	1
Primary B	*hidgish*, 'emphasize'	1
Primary B	*hudlaf*, 'be leaked'	1
Primary B	*huvhar*, 'be clarified'	1
Primary B	*maca*, 'found'	1
Primary B	*ramaz*, 'hint, indicate'	1
Primary B	*rimez*, 'hint, indicate'	1
Primary B	*yivada*, 'become known'	1
Secondary A	*hispik*, 'be sufficient'	20
Secondary A	*hifsik*, 'stop'	3
Secondary A	*hizdamen*, 'happen arrive'	3
Secondary A	*hicliax*, 'succeed'	1
Secondary A	*tixnen*, 'plan'	1
Secondary B	*heemin*, 'believe'	88
Secondary B	*higia*, 'deserve'	80
Secondary B	*ixpat*, 'care'	53
Secondary B	*xika*, 'wait'	26
Secondary B	*higia*, 'arrive'	14
Secondary B	*ba*, 'want'	12
Secondary B	*himlic*, 'recommend'	10
Secondary B	*nimas*, 'be fed up with'	10
Secondary B	*huca*, 'be suggested'	7
Secondary B	*ba*, 'come'	6
Secondary B	*hitxashek*, 'feel like'	5
Secondary B	*xara*, 'upset'	5
Secondary B	*himtin*, 'wait'	4
Secondary B	*zakuk*, 'in need of'	4
Secondary B	*cafuy*, 'predictable'	3
Secondary B	*cipa*, 'expect'	2
Secondary B	*hitxayev*, 'be obligated'	2
Secondary B	*nishba*, 'swear'	2
Secondary B	*arev*, 'pleasant'	1
Secondary B	*racuy*, 'wanted'	1
Secondary C	*azar*, 'help'	250
Secondary C	*ifsher*, 'enable'	226

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
Secondary C	*asa*, 'make'	135
Secondary C	*siyea*, 'aid'	63
Secondary C	*hirsha*, 'allow'	58
Secondary C	*mutar*, 'allowed'	57
Secondary C	*ixel*, 'wish'	49
Secondary C	*garam*, 'cause'	45
Secondary C	*isher*, 'approve'	33
Secondary C	*daag*, 'worry'	17
Secondary C	*hitir*, 'allow'	15
Secondary C	*hursha*, 'be allowed'	10
Secondary C	*nigram*, 'be caused'	8
Secondary C	*viter*, 'concede'	7
Secondary C	*daag*, 'take care of'	6
Secondary C	*ushar*, 'be approved'	6
Secondary C	*hitafsher*, 'be enabled'	5
Secondary C	*hora*, 'instruct'	5
Secondary C	*hixtiv*, 'dictate'	4
Secondary C	*amar*, 'say'	1
Secondary C	*huxtav*, 'be dictated'	1
Secondary C	*nidrash*, 'be needed'	1
Secondary C	*nifsak*, 'be ruled'	1
Secondary D	*nire*, 'seem'	284
Secondary D	*nidme*, 'seem'	230
Secondary D	*nishma*, 'be heard'	14
Secondary D	*meshane*, 'matter'	10
Secondary D	*asa*, 'make'	1
PROPERTY	*barur*, 'clear'	137
PROPERTY	*kashe*, 'hard'	112
PROPERTY	*xaser*, 'lacking'	71
PROPERTY	*noax*, 'comfortable'	38
PROPERTY	*hitim*, 'fit, suit'	35
PROPERTY	*car*, 'narrow'	32
PROPERTY	*koev*, 'painful'	20
PROPERTY	*notar*, 'be left'	19
PROPERTY	*naim*, 'pleasant'	17
PROPERTY	*kal*, 'easy'	13

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
PROPERTY	*histalem*, 'be worthwhile'	12
PROPERTY	*kafuf*, 'subordinate'	10
PROPERTY	*xadash*, 'new'	7
PROPERTY	*corem*, 'irritating, bothering'	6
PROPERTY	*nagish*, 'accessible'	6
PROPERTY	*rauy*, 'deserve'	5
PROPERTY	*kodem*, 'prior'	2
PROPERTY	*meshutaf*, 'common'	2
PROPERTY	*acuv*, 'sad'	1
PROPERTY	*atum*, 'opaque'	1
PROPERTY	*ba'ar*, 'burn'	1
PROPERTY	*kadam*, 'prior'	1
PROPERTY	*kashur*, 'related'	1
PROPERTY	*masriax*, 'smelly'	1
PROPERTY	*meanyen*, 'interesting'	1
PROPERTY	*ofyani*, 'characteristic'	1
PROPERTY	*ravax*, 'be relieved'	1
PROPERTY	*shamur*, 'reserved'	1
VALUE	*xashuv*, 'important'	181
VALUE	*asur*, 'forbidden'	79
VALUE	*hifria*, 'disturb'	60
VALUE	*tov*, 'good'	27
VALUE	*keday*, 'worthwhile'	25
VALUE	*hecik*, 'bother'	13
VALUE	*xaval*, 'pity'	11
VALUE	*shave*, 'equal'	8
VALUE	*adif*, 'preferable'	7
VALUE	*naxuc*, 'required'	6
VALUE	*yakar*, 'expensive'	6
VALUE	*naxon*, 'correct'	4
VALUE	*ra*, 'bad'	3
VALUE	*relevanti*, 'relevant'	3
VALUE	*mesukan*, 'dangerous'	2
VALUE	*meyuxad*, 'special'	2
VALUE	*muzar*, 'strange'	2
VALUE	*nidned*, 'nag'	2
VALUE	*shve erex*, 'equivalent'	2

Continued on next page

Table A.1 – *Continued from previous page*

Type	Predicate	Frequency
VALUE	*heik*, 'bother'	1
VALUE	*hitamer*, 'abuse'	1
VALUE	*hitnakel*, 'harass'	1
VALUE	*kriti*, 'critic'	1
VALUE	*mehave hanxaya*, 'be guidance'	1
VALUE	*meorer hashraa*, 'inspire'	1
VALUE	*nifla*, 'wonderful'	1
VALUE	*ragil*, 'regular'	1
VALUE	*trivyali*, 'trivial'	1
VALUE	*zol*, 'cheap'	1
IRR	*toda*, 'thanks'	82
IRR	*kavod*, 'respect'	20
IRR	*lama*, 'why'	10
IRR	*al*, 'don't'	4
IRR	*day*, 'enough'	4
IRR	*mixuc*, 'out of'	4
IRR	*oy*, 'oh my'	4
IRR	*boker tov*, 'good morning'	3
IRR	*kef*, 'fun'	2
IRR	*minayn*, 'from whence'	2
IRR	*shalom*, 'hello'	2
IRR	*sheela*, 'question'	2
IRR	*shutaf*, 'partner'	2
IRR	*avel*, 'wrong'	1
IRR	*barax*, 'escape'	1
IRR	*bishvil ma*, 'for what'	1
IRR	*braxot*, 'congratulations'	1
IRR	*havana*, 'understanding'	1
IRR	*leoneg*, 'pleasure'	1
IRR	*misaviv*, 'around'	1
IRR	*pesha*, '(a) crime'	1
IRR	*samux*, 'adjacent'	1
IRR	*tshuva*, 'answer'	1
IRR	*xalila*, 'God forbid!'	1
IRR	*xashash*, 'fear'	1
IRR	*xova*, 'requirement'	1
IRR	*yeshar koax*, 'well done!'	1

Table A.2: Frequency of predicates according to Dative function

Dative function	**Predicate**	**Frequency**	**% of Dative function**
ADD	*amar*, 'say'	1,436	43.81%
ADD	*higid*, 'tell'	532	16.23%
ADD	*siper*, 'tell'	141	4.30%
ADD	*hicia*, 'suggest'	115	3.51%
ADD	*hisbir*, 'explain'	113	3.45%
ADD	*hera*, 'show'	101	3.08%
ADD	*ana*, 'answer'	86	2.62%
ADD	*hodia*, 'inform'	84	2.56%
ADD	*hivtiax*, 'promise'	64	1.95%
ADD	*hicig*, 'present'	58	1.77%
ADD	*hizkir*, 'remind'	51	1.56%
ADD	*ixel*, 'wish'	49	1.49%
ADD	*kara*, 'call'	46	1.40%
ADD	*katav*, 'write'	39	1.19%
ADD	*nimsar*, 'be delivered'	30	0.92%
ADD	*diveax*, 'report'	24	0.73%
ADD	*masar*, 'deliver'	24	0.73%
ADD	*hivhir*, 'clarify'	19	0.58%
ADD	*gila*, 'discover'	16	0.49%
ADD	*hikri*, 'read aloud'	14	0.43%
ADD	*heshiv*, 'return, answer, reply'	12	0.37%
ADD	*hoxiax*, 'prove'	12	0.37%
ADD	*kara*, 'read'	12	0.37%
ADD	*hucag*, 'be presented'	11	0.34%
ADD	*himlic*, 'recommend'	10	0.31%
ADD	*teer*, 'describe'	10	0.31%
ADD	*rasham*, 'write, note'	9	0.27%
ADD	*biser*, 'inform'	8	0.24%
ADD	*heir*, 'remark'	8	0.24%
ADD	*himxish*, 'illustrate'	7	0.21%
ADD	*huca*, 'be suggested'	7	0.21%
ADD	*peret*, 'describe, detail, elaborate'	6	0.18%
ADD	*xidesh*, 'Re-establish, renewed'	6	0.18%
ADD	*caak*, 'shout'	5	0.15%
ADD	*hora*, 'instruct'	5	0.15%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
ADD	*isher*, 'approve'	5	0.15%
ADD	*rashum*, 'written'	5	0.15%
ADD	*xatam*, 'sign'	5	0.15%
ADD	*hexmi*, 'flatter'	4	0.12%
ADD	*hixtiv*, 'dictate'	4	0.12%
ADD	*katuv*, 'written'	4	0.12%
ADD	*asa*, 'make'	3	0.09%
ADD	*boker tov*, 'good morning'	3	0.09%
ADD	*citet*, 'quote'	3	0.09%
ADD	*ciyen*, 'note'	3	0.09%
ADD	*higdir*, 'define'	3	0.09%
ADD	*hishmia*, 'make a sound'	3	0.09%
ADD	*hugdar*, 'be defined'	3	0.09%
ADD	*kara*, 'name'	3	0.09%
ADD	*laag*, 'mock'	3	0.09%
ADD	*tirgem*, 'translate'	3	0.09%
ADD	*zarak*, 'throw'	3	0.09%
ADD	*duvax*, 'be reported'	2	0.06%
ADD	*hidgim*, 'demonstrate'	2	0.06%
ADD	*hikrin*, 'project'	2	0.06%
ADD	*hitif*, 'preach'	2	0.06%
ADD	*katuv*, 'writen'	2	0.06%
ADD	*nidned*, 'nag'	2	0.06%
ADD	*nishba*, 'swear'	2	0.06%
ADD	*ba-beteana*, 'complain'	1	0.03%
ADD	*ba*, 'come'	1	0.03%
ADD	*ciyer*, 'paint'	1	0.03%
ADD	*hecif*, 'flood'	1	0.03%
ADD	*hicbia*, 'vote'	1	0.03%
ADD	*hidgish*, 'emphasize'	1	0.03%
ADD	*hifna*, 'turn'	1	0.03%
ADD	*hikir*, 'introduce'	1	0.03%
ADD	*hishlim*, 'complete'	1	0.03%
ADD	*hitgala*, 'be discovered'	1	0.03%
ADD	*hitna*, 'condition'	1	0.03%
ADD	*hitva*, 'outline'	1	0.03%
ADD	*hokir*, 'cherish'	1	0.03%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
ADD	*hudlaf*, 'be leaked'	1	0.03%
ADD	*huvhar*, 'be clarified'	1	0.03%
ADD	*huxtav*, 'be dictated'	1	0.03%
ADD	*karac*, 'wink'	1	0.03%
ADD	*laxash*, 'whisper'	1	0.03%
ADD	*mana*, 'count'	1	0.03%
ADD	*mehave hanxaya*, 'be guidance'	1	0.03%
ADD	*natan*, 'give'	1	0.03%
ADD	*nifnef*, 'wave'	1	0.03%
ADD	*nixtav*, 'be writen'	1	0.03%
ADD	*nixtav*, 'written'	1	0.03%
ADD	*pina*, 'evacuate'	1	0.03%
ADD	*ramaz*, 'hint, indicate'	1	0.03%
ADD	*rimez*, 'hint, indicate'	1	0.03%
ADD	*shikef*, 'reflect'	1	0.03%
ADD	*shiker*, 'lie'	1	0.03%
ADD	*siber*, 'give a simple explanation'	1	0.03%
ADD	*tava*, 'drown'	1	0.03%
ADD	*tiftef*, 'drip'	1	0.03%
ADD	*tshuva*, 'answer'	1	0.03%
ADD	*ushar*, 'be approved'	1	0.03%
AFF	*hexin*, 'prepare'	25	15.72%
AFF	*asa*, 'make'	22	13.84%
AFF	*patax*, 'open'	11	6.92%
AFF	*bana*, 'build'	10	6.29%
AFF	*kana*, 'buy'	9	5.66%
AFF	*maca*, 'find'	8	5.03%
AFF	*sider*, 'organize, fix'	7	4.40%
AFF	*hoci*, 'take out'	6	3.77%
AFF	*xipes*, 'search'	6	3.77%
AFF	*day*, 'enough'	4	2.52%
AFF	*badak*, 'check'	3	1.89%
AFF	*barax*, 'escape'	3	1.89%
AFF	*nimca*, 'be found'	3	1.89%
AFF	*patar*, 'solve'	3	1.89%
AFF	*heela*, 'raise'	2	1.26%
AFF	*heemid*, 'position'	2	1.26%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
AFF	*ikev*, 'delay'	2	1.26%
AFF	*mile*, 'fill'	2	1.26%
AFF	*shixrer*, 'release'	2	1.26%
AFF	*tipel*, 'take care of'	2	1.26%
AFF	*acar*, 'stop'	1	0.63%
AFF	*asaf*, 'gather'	1	0.63%
AFF	*baxan*, 'examine'	1	0.63%
AFF	*baxar*, 'choose'	1	0.63%
AFF	*bishvil ma*, 'for what'	1	0.63%
AFF	*hefic*, 'distribute'	1	0.63%
AFF	*heir*, 'illuminate'	1	0.63%
AFF	*hekim*, 'establish'	1	0.63%
AFF	*higdil*, 'cause to grow'	1	0.63%
AFF	*hikpi*, 'freeze'	1	0.63%
AFF	*hisig*, 'accomplish'	1	0.63%
AFF	*hitir*, 'allow'	1	0.63%
AFF	*huxan*, 'be prepared'	1	0.63%
AFF	*ifyen*, 'characterize'	1	0.63%
AFF	*irgen*, 'organize'	1	0.63%
AFF	*kisa*, 'cover'	1	0.63%
AFF	*niftax*, 'be opened'	1	0.63%
AFF	*nivna*, 'be built'	1	0.63%
AFF	*rikez*, 'assemble'	1	0.63%
AFF	*saxar*, 'hire'	1	0.63%
AFF	*shamar*, 'save'	1	0.63%
AFF	*shamur*, 'reserved'	1	0.63%
AFF	*shiklel*, 'weigh'	1	0.63%
AFF	*sibsed*, 'subsidize'	1	0.63%
AFF	*sikem*, 'summarize'	1	0.63%
AFF	*xakar*, 'investigate'	1	0.63%
AFF	*xilek*, 'divide, hand out'	1	0.63%
DM	*hoda*, 'thank'	82	27.06%
DM	*heemin*, 'believe'	80	26.40%
DM	*nidme*, 'seem'	47	15.51%
DM	*salax*, 'forgive'	41	13.53%
DM	*kavod*, 'respect'	20	6.60%
DM	*da*, 'know'	18	5.94%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
DM	*hursha*, 'be allowed'	10	3.30%
DM	*mutar*, 'allowed'	2	0.66%
DM	*shalom*, 'hello'	2	0.66%
DM	*yeshar koax*, 'well done!'	1	0.33%
ETH	*asa*, 'make'	6	12.24%
ETH	*kafac*, 'jump'	4	8.16%
ETH	*nixnas*, 'enter'	4	8.16%
ETH	*hicmid*, 'attach'	3	6.12%
ETH	*yashav*, 'sit'	3	6.12%
ETH	*amad*, 'stand'	2	4.08%
ETH	*hofia*, 'appear'	2	4.08%
ETH	*horid*, 'reduce'	2	4.08%
ETH	*kadax*, 'drill'	2	4.08%
ETH	*nafal*, 'fall'	2	4.08%
ETH	*neelam*, 'disappear'	2	4.08%
ETH	*safar*, 'count'	2	4.08%
ETH	*xazar*, 'return'	2	4.08%
ETH	*baxa*, 'cry'	1	2.04%
ETH	*gadal*, 'grow'	1	2.04%
ETH	*hictanea*, 'be modest'	1	2.04%
ETH	*hishpia*, 'influence'	1	2.04%
ETH	*hishtolel*, 'be unruly'	1	2.04%
ETH	*hithapex*, 'turn over'	1	2.04%
ETH	*kam*, 'get up'	1	2.04%
ETH	*kava*, 'determine'	1	2.04%
ETH	*neecar*, 'be delayed'	1	2.04%
ETH	*niftar*, 'be solved'	1	2.04%
ETH	*nihel*, 'manage'	1	2.04%
ETH	*shatal*, 'plant'	1	2.04%
ETH	*tixnen*, 'plan'	1	2.04%
EVAL	*xashuv*, 'important'	181	44.15%
EVAL	*asur*, 'forbidden'	79	19.27%
EVAL	*mutar*, 'allowed'	54	13.17%
EVAL	*tov*, 'good'	27	6.59%
EVAL	*keday*, 'worthwhile'	25	6.10%
EVAL	*lama*, 'why'	10	2.44%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
EVAL	*adif*, 'preferable'	7	1.71%
EVAL	*yakar*, 'expensive'	6	1.46%
EVAL	*rauy*, 'deserve'	5	1.22%
EVAL	*al*, 'don't'	4	0.98%
EVAL	*naxon*, 'correct'	4	0.98%
EVAL	*relevanti*, 'relevant'	3	0.73%
EVAL	*mesukan*, 'dangerous'	2	0.49%
EVAL	*hitim*, 'fit, suit'	1	0.24%
EVAL	*ofyani*, 'characteristic'	1	0.24%
EVAL	*racuy*, 'wanted'	1	0.24%
EXP	*nire*, 'seem'	284	15.56%
EXP	*nidme*, 'seem'	183	10.03%
EXP	*barur*, 'clear'	137	7.51%
EXP	*kashe*, 'hard'	112	6.14%
EXP	*yadua*, 'known'	94	5.15%
EXP	*higia*, 'deserve'	79	4.33%
EXP	*xaser*, 'lacking'	71	3.89%
EXP	*karah*, 'happen'	67	3.67%
EXP	*hifria*, 'disturb'	60	3.29%
EXP	*ixpat*, 'care'	53	2.90%
EXP	*ala*, 'rise, cost'	47	2.58%
EXP	*lakax*, 'take'	47	2.58%
EXP	*noax*, 'comfortable'	38	2.08%
EXP	*car*, 'narrow'	32	1.75%
EXP	*hitim*, 'fit, suit'	32	1.75%
EXP	*yaca*, 'got to do'	26	1.42%
EXP	*nishar*, 'be left'	22	1.21%
EXP	*hispik*, 'be sufficient'	20	1.10%
EXP	*koev*, 'painful'	20	1.10%
EXP	*notar*, 'be left'	19	1.04%
EXP	*mukar*, 'known'	18	0.99%
EXP	*naim*, 'pleasant'	17	0.93%
EXP	*histaber*, 'become clear'	16	0.88%
EXP	*hitbarer*, 'become clear'	16	0.88%
EXP	*noda*, 'become known'	16	0.88%
EXP	*teer*, 'describe'	15	0.82%
EXP	*higia*, 'arrive'	14	0.77%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
EXP	*nishma*, 'be heard'	14	0.77%
EXP	*hecik*, 'bother'	13	0.71%
EXP	*kal*, 'easy'	13	0.71%
EXP	*ba*, 'want'	12	0.66%
EXP	*zaxur*, 'be remembered'	12	0.66%
EXP	*xaval*, 'pity'	11	0.60%
EXP	*meshane*, 'matter'	10	0.55%
EXP	*nimas*, 'be fed up with'	10	0.55%
EXP	*nogea*, 'concern'	8	0.44%
EXP	*shave*, 'equal'	8	0.44%
EXP	*mal'u*, 'to reach a certain age'	7	0.38%
EXP	*viter*, 'concede'	7	0.38%
EXP	*xadash*, 'new'	7	0.38%
EXP	*corem*, 'irritating, bothering'	6	0.33%
EXP	*histader*, 'work out'	6	0.33%
EXP	*nagish*, 'accessible'	6	0.33%
EXP	*naxuc*, 'required'	6	0.33%
EXP	*ba*, 'come'	5	0.27%
EXP	*hitxashek*, 'feel like'	5	0.27%
EXP	*muvan*, 'understood'	5	0.27%
EXP	*xara*, 'upset'	5	0.27%
EXP	*nishkaf*, 'be foreseen'	4	0.22%
EXP	*oy*, 'oh my'	4	0.22%
EXP	*cafuy*, 'predictable'	3	0.16%
EXP	*darush*, 'be needed'	3	0.16%
EXP	*halax*, 'go'	3	0.16%
EXP	*hitafsher*, 'be enabled'	3	0.16%
EXP	*hitarev*, 'interrupt'	3	0.16%
EXP	*hitvasef*, 'be added'	3	0.16%
EXP	*hizdamen*, 'happen arrive'	3	0.16%
EXP	*ra*, 'bad'	3	0.16%
EXP	*yaca*, 'exit'	3	0.16%
EXP	*avar*, 'pass'	2	0.11%
EXP	*bilbel*, 'confuse'	2	0.11%
EXP	*heniax*, 'leave alone, set down'	2	0.11%
EXP	*kef*, 'fun'	2	0.11%
EXP	*meyuxad*, 'special'	2	0.11%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
EXP	*minayn*, 'from whence'	2	0.11%
EXP	*muzar*, 'strange'	2	0.11%
EXP	*nocar*, 'be created'	2	0.11%
EXP	*sheela*, 'question'	2	0.11%
EXP	*zar*, 'strange'	2	0.11%
EXP	*acuv*, 'sad'	1	0.05%
EXP	*asa*, 'make'	1	0.05%
EXP	*atum*, 'opaque'	1	0.05%
EXP	*ba'ar*, 'burn'	1	0.05%
EXP	*heela*, 'raise'	1	0.05%
EXP	*heik*, 'bother'	1	0.05%
EXP	*her'a*, 'show'	1	0.05%
EXP	*heziz*, 'move something'	1	0.05%
EXP	*hicliax*, 'succeed'	1	0.05%
EXP	*hisvev*, 'turn around'	1	0.05%
EXP	*hitamer*, 'abuse'	1	0.05%
EXP	*hitgala*, 'be discovered'	1	0.05%
EXP	*hitnakel*, 'harass'	1	0.05%
EXP	*hitparec*, 'burst in'	1	0.05%
EXP	*hitpocec*, 'explode'	1	0.05%
EXP	*hitxaber*, 'become attached'	1	0.05%
EXP	*hivrik*, 'shine'	1	0.05%
EXP	*hix'iv*, 'hurt'	1	0.05%
EXP	*hutal*, 'be imposed'	1	0.05%
EXP	*kasam*, 'enchant'	1	0.05%
EXP	*katua*, 'be disrupted'	1	0.05%
EXP	*kriti*, 'critic'	1	0.05%
EXP	*laxuc*, 'pressed, stressed'	1	0.05%
EXP	*leoneg*, 'pleasure'	1	0.05%
EXP	*masriax*, 'smelly'	1	0.05%
EXP	*meanyen*, 'interesting'	1	0.05%
EXP	*meorer hashraa*, 'inspire'	1	0.05%
EXP	*nidlak*, 'lit, to be lighted'	1	0.05%
EXP	*nidrash*, 'be needed'	1	0.05%
EXP	*nifla*, 'wonderful'	1	0.05%
EXP	*nigmar*, 'over'	1	0.05%
EXP	*nikba*, 'be decided'	1	0.05%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
EXP	*ravax*, 'be relieved'	1	0.05%
EXP	*trivyali*, 'trivial'	1	0.05%
EXP	*xayav*, 'owe'	1	0.05%
EXP	*xazar*, 'return'	1	0.05%
EXP	*xova*, 'requirement'	1	0.05%
EXP	*yivada*, 'become known'	1	0.05%
EXP	*zol*, 'cheap'	1	0.05%
HEP	*hoda*, 'thank'	277	15.35%
HEP	*azar*, 'help'	250	13.86%
HEP	*ifsher*, 'enable'	226	12.53%
HEP	*natan*, 'give'	178	9.87%
HEP	*kara*, 'name'	128	7.10%
HEP	*asa*, 'make'	102	5.65%
HEP	*siyea*, 'aid'	63	3.49%
HEP	*hirsha*, 'allow'	58	3.22%
HEP	*garam*, 'cause'	45	2.49%
HEP	*isher*, 'approve'	28	1.55%
HEP	*xika*, 'wait'	26	1.44%
HEP	*hikshiv*, 'listen'	20	1.11%
HEP	*daag*, 'worry'	17	0.94%
HEP	*kava*, 'determine'	17	0.94%
HEP	*yicer*, 'produce'	15	0.83%
HEP	*hitir*, 'allow'	14	0.78%
HEP	*maca*, 'find'	13	0.72%
HEP	*histalem*, 'be worthwhile'	12	0.67%
HEP	*salax*, 'forgive'	12	0.67%
HEP	*nikba*, 'be decided'	11	0.61%
HEP	*patax*, 'open'	11	0.61%
HEP	*hikdish*, 'dedicate'	10	0.55%
HEP	*kafuf*, 'subordinate'	10	0.55%
HEP	*heemin*, 'believe'	8	0.44%
HEP	*nigram*, 'be caused'	8	0.44%
HEP	*taram*, 'donate'	8	0.44%
HEP	*yixes*, 'relate'	8	0.44%
HEP	*hoci*, 'take out'	7	0.39%
HEP	*xasax*, 'save'	7	0.39%
HEP	*daag*, 'take care of'	6	0.33%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
HEP	*hoil*, 'be useful'	6	0.33%
HEP	*nimca*, 'be found'	6	0.33%
HEP	*xayav*, 'owe'	6	0.33%
HEP	*arax*, 'organize'	5	0.28%
HEP	*hirbic*, 'hit'	5	0.28%
HEP	*shamar*, 'save'	5	0.28%
HEP	*shimesh*, 'be used'	5	0.28%
HEP	*ushar*, 'be approved'	5	0.28%
HEP	*hidbik*, 'glue'	4	0.22%
HEP	*himtin*, 'wait'	4	0.22%
HEP	*hukca*, 'be allocated'	4	0.22%
HEP	*lakax*, 'take'	4	0.22%
HEP	*mixuc*, 'out of'	4	0.22%
HEP	*naana*, 'be accepted'	4	0.22%
HEP	*niftax*, 'be opened'	4	0.22%
HEP	*shama*, 'hear'	4	0.22%
HEP	*yuxas*, 'be related'	4	0.22%
HEP	*zakuk*, 'in need of'	4	0.22%
HEP	*arev*, 'guarantee'	3	0.17%
HEP	*hikca*, 'allocate'	3	0.17%
HEP	*hizik*, 'harm'	3	0.17%
HEP	*horid*, 'reduce'	3	0.17%
HEP	*raxash*, 'feel (respect)'	3	0.17%
HEP	*sam*, 'put'	3	0.17%
HEP	*xipes*, 'search'	3	0.17%
HEP	*cipa*, 'expect'	2	0.11%
HEP	*heezin*, 'listen attentively'	2	0.11%
HEP	*hegiv*, 'respond'	2	0.11%
HEP	*hexlif*, 'exchange'	2	0.11%
HEP	*hicmid*, 'attach'	2	0.11%
HEP	*hikir*, 'acknowledge'	2	0.11%
HEP	*hishlim*, 'complete'	2	0.11%
HEP	*hitafsher*, 'be enabled'	2	0.11%
HEP	*hitim*, 'fit, suit'	2	0.11%
HEP	*hitxayev*, 'be obligated'	2	0.11%
HEP	*hukdash*, 'be dedicated'	2	0.11%
HEP	*kodem*, 'prior'	2	0.11%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
HEP	*meshutaf*, 'common'	2	0.11%
HEP	*na'asa*, 'be made'	2	0.11%
HEP	*neetar*, 'accede'	2	0.11%
HEP	*nitek*, 'disconnect'	2	0.11%
HEP	*nolad*, 'born'	2	0.11%
HEP	*shar*, 'sing'	2	0.11%
HEP	*shutaf*, 'partner'	2	0.11%
HEP	*shve erex*, 'equivalent'	2	0.11%
HEP	*sider*, 'organize, fix'	2	0.11%
HEP	*xared*, 'scared'	2	0.11%
HEP	*xata*, 'sin'	2	0.11%
HEP	*arev*, 'pleasant'	1	0.06%
HEP	*bana*, 'build'	1	0.06%
HEP	*benigud*, 'as opposed to'	1	0.06%
HEP	*ceref*, 'add'	1	0.06%
HEP	*cilem*, 'photograph'	1	0.06%
HEP	*dafak*, 'knock, bang'	1	0.06%
HEP	*gamal*, 'recompense'	1	0.06%
HEP	*gilgel*, 'roll'	1	0.06%
HEP	*havana*, 'understanding'	1	0.06%
HEP	*heir*, 'illuminate'	1	0.06%
HEP	*hevi*, 'bring'	1	0.06%
HEP	*hexdir*, 'insert'	1	0.06%
HEP	*hicdia*, 'salute'	1	0.06%
HEP	*hictaber*, 'accumulate'	1	0.06%
HEP	*hikciv*, 'allocate'	1	0.06%
HEP	*hikri*, 'read aloud'	1	0.06%
HEP	*hishir*, 'leave'	1	0.06%
HEP	*hitnaged*, 'resist'	1	0.06%
HEP	*hitnaker*, 'ignore, alienate'	1	0.06%
HEP	*hitxaxesh*, 'deny'	1	0.06%
HEP	*hixnis*, 'insert'	1	0.06%
HEP	*hizrim*, 'cause flow'	1	0.06%
HEP	*hudbak*, 'be glued'	1	0.06%
HEP	*hurad*, 'be reduced'	1	0.06%
HEP	*kacav*, 'allocate'	1	0.06%
HEP	*kadam*, 'prior'	1	0.06%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
HEP	*kara*, 'read'	1	0.06%
HEP	*karav*, 'approach'	1	0.06%
HEP	*kashur*, 'related'	1	0.06%
HEP	*kicec*, 'cut off'	1	0.06%
HEP	*maca*, 'found'	1	0.06%
HEP	*maxal*, 'forgive'	1	0.06%
HEP	*mecoraf*, 'attached'	1	0.06%
HEP	*misaviv*, 'around'	1	0.06%
HEP	*mutar*, 'allowed'	1	0.06%
HEP	*ne'erax*, 'be organized'	1	0.06%
HEP	*nicbar*, 'be collected'	1	0.06%
HEP	*nidbak*, 'be glued'	1	0.06%
HEP	*nifsak*, 'be ruled'	1	0.06%
HEP	*nigzar*, 'be sentenced'	1	0.06%
HEP	*nika*, 'deduct'	1	0.06%
HEP	*nizkak*, 'need'	1	0.06%
HEP	*nocar*, 'be created'	1	0.06%
HEP	*pasak*, 'rule'	1	0.06%
HEP	*pesha*, '(a) crime'	1	0.06%
HEP	*pina*, 'evacuate'	1	0.06%
HEP	*prat*, 'excluding'	1	0.06%
HEP	*ragil*, 'regular'	1	0.06%
HEP	*samux*, 'adjacent'	1	0.06%
HEP	*shina*, 'change, matter'	1	0.06%
HEP	*shiyex*, 'attribute'	1	0.06%
HEP	*shudax*, 'be paired'	1	0.06%
HEP	*tikcev*, 'allocate'	1	0.06%
HEP	*tiken*, 'fix, amend'	1	0.06%
HEP	*xalila*, 'God forbid!'	1	0.06%
HEP	*xashash*, 'fear'	1	0.06%
HEP	*xiber*, 'join'	1	0.06%
HEP	*yied*, 'designate'	1	0.06%
HEP	*yixed*, 'dedicate'	1	0.06%
POSSDAT	*lakax*, 'take'	27	12.05%
POSSDAT	*haras*, 'destroy'	12	5.36%
POSSDAT	*horid*, 'reduce'	12	5.36%
POSSDAT	*patar*, 'solve'	8	3.57%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
POSSDAT	*dafak*, 'knock, bang'	5	2.23%
POSSDAT	*sagar*, 'close'	5	2.23%
POSSDAT	*shalal*, 'rule out'	5	2.23%
POSSDAT	*bitel*, 'cancel'	4	1.79%
POSSDAT	*ganav*, 'steal'	4	1.79%
POSSDAT	*hixpil*, 'multiply'	4	1.79%
POSSDAT	*nitek*, 'disconnect'	4	1.79%
POSSDAT	*shina*, 'change, matter'	4	1.79%
POSSDAT	*yarad*, 'descend'	4	1.79%
POSSDAT	*acar*, 'stop'	3	1.34%
POSSDAT	*amad*, 'stand'	3	1.34%
POSSDAT	*gamar*, 'finish'	3	1.34%
POSSDAT	*hafax*, 'turn over'	3	1.34%
POSSDAT	*hifsik*, 'stop'	3	1.34%
POSSDAT	*kara'*, 'tear'	3	1.34%
POSSDAT	*nihel*, 'manage'	3	1.34%
POSSDAT	*patax*, 'open'	3	1.34%
POSSDAT	*pocec*, 'destroy'	3	1.34%
POSSDAT	*yiker*, 'raise the price'	3	1.34%
POSSDAT	*af*, 'fly'	2	0.89%
POSSDAT	*asa*, 'make'	2	0.89%
POSSDAT	*avad*, 'got lost'	2	0.89%
POSSDAT	*barax*, 'escape'	2	0.89%
POSSDAT	*daxa*, 'postpone'	2	0.89%
POSSDAT	*hexlif*, 'exchange'	2	0.89%
POSSDAT	*hishlim*, 'complete'	2	0.89%
POSSDAT	*hishtalet*, 'gain control'	2	0.89%
POSSDAT	*histakel*, 'look'	2	0.89%
POSSDAT	*hitpocec*, 'explode'	2	0.89%
POSSDAT	*maca*, 'find'	2	0.89%
POSSDAT	*nafal*, 'fall'	2	0.89%
POSSDAT	*paga*, 'harm'	2	0.89%
POSSDAT	*palash*, 'invade'	2	0.89%
POSSDAT	*saraf*, 'burn'	2	0.89%
POSSDAT	*shata*, 'drink'	2	0.89%
POSSDAT	*yashav*, 'sit'	2	0.89%
POSSDAT	*avel*, 'wrong'	1	0.45%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
POSSDAT	*badak*, 'check'	1	0.45%
POSSDAT	*bana*, 'build'	1	0.45%
POSSDAT	*cimcem*, 'reduce'	1	0.45%
POSSDAT	*ciyer*, 'paint'	1	0.45%
POSSDAT	*dalaf*, 'leak'	1	0.45%
POSSDAT	*gilgel*, 'roll'	1	0.45%
POSSDAT	*halax*, 'go'	1	0.45%
POSSDAT	*harag*, 'kill'	1	0.45%
POSSDAT	*heela*, 'raise'	1	0.45%
POSSDAT	*heelim*, 'conceal'	1	0.45%
POSSDAT	*heerix*, 'lengthen, extend '	1	0.45%
POSSDAT	*herim*, 'lift'	1	0.45%
POSSDAT	*hexrim*, 'confiscate'	1	0.45%
POSSDAT	*hexshiv*, 'ascribe inportance'	1	0.45%
POSSDAT	*hifxit*, 'subtract'	1	0.45%
POSSDAT	*hilbin*, 'whiten, bleach'	1	0.45%
POSSDAT	*hipil*, 'drop'	1	0.45%
POSSDAT	*hir'il*, 'poison'	1	0.45%
POSSDAT	*hishxit*, 'vandalize'	1	0.45%
POSSDAT	*hitxalef*, 'be changed'	1	0.45%
POSSDAT	*hofia*, 'appear'	1	0.45%
POSSDAT	*ir'er*, 'undercut, destabilize'	1	0.45%
POSSDAT	*kafac*, 'jump'	1	0.45%
POSSDAT	*karat*, 'cut down'	1	0.45%
POSSDAT	*kata*, 'interrupt'	1	0.45%
POSSDAT	*kava*, 'determine'	1	0.45%
POSSDAT	*kicec*, 'cut off'	1	0.45%
POSSDAT	*kicer*, 'shorten'	1	0.45%
POSSDAT	*kilkel*, 'damage'	1	0.45%
POSSDAT	*kiseax*, 'hit violently'	1	0.45%
POSSDAT	*kizez*, 'deduct'	1	0.45%
POSSDAT	*maxak*, 'erase'	1	0.45%
POSSDAT	*mile*, 'fill'	1	0.45%
POSSDAT	*misa*, 'tax'	1	0.45%
POSSDAT	*naga*, 'touch'	1	0.45%
POSSDAT	*neheras*, 'be destroyed'	1	0.45%
POSSDAT	*nexsam*, 'be blocked'	1	0.45%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
POSSDAT	*nifga*, 'be injured'	1	0.45%
POSSDAT	*niftax*, 'be opened'	1	0.45%
POSSDAT	*nikev*, 'puncture'	1	0.45%
POSSDAT	*nilkax*, 'be taken'	1	0.45%
POSSDAT	*nisgar*, 'be closed'	1	0.45%
POSSDAT	*nitpas*, 'be caught'	1	0.45%
POSSDAT	*parac*, 'break in'	1	0.45%
POSSDAT	*perek*, 'dissemble'	1	0.45%
POSSDAT	*pina*, 'evacuate'	1	0.45%
POSSDAT	*pizer*, 'scatter'	1	0.45%
POSSDAT	*roked*, 'dance'	1	0.45%
POSSDAT	*satam*, 'seal'	1	0.45%
POSSDAT	*shafax*, 'spill'	1	0.45%
POSSDAT	*shamat*, 'drop'	1	0.45%
POSSDAT	*shavar*, 'break'	1	0.45%
POSSDAT	*shibesh*, 'disturb'	1	0.45%
POSSDAT	*shidreg*, 'upgrade'	1	0.45%
POSSDAT	*tafas*, 'catch'	1	0.45%
POSSDAT	*talash*, 'detach'	1	0.45%
POSSDAT	*utar*, 'be located'	1	0.45%
POSSDAT	*xarat*, 'engrave'	1	0.45%
POSSDAT	*xatam*, 'sign'	1	0.45%
POSSDAT	*xazar*, 'return'	1	0.45%
POSSDAT	*zarak*, 'throw'	1	0.45%
REC	*natan*, 'give'	1,025	62.42%
REC	*heevir*, 'transfer'	94	5.72%
REC	*hevi*, 'bring'	79	4.81%
REC	*shilem*, 'pay'	74	4.51%
REC	*nitan*, 'be given'	57	3.47%
REC	*hexzir*, 'return'	45	2.74%
REC	*shalax*, 'send'	44	2.68%
REC	*sipek*, 'supply'	32	1.95%
REC	*hosif*, 'add'	28	1.71%
REC	*higish*, 'serve'	25	1.52%
REC	*hishir*, 'leave'	20	1.22%
REC	*maxar*, 'sell'	12	0.73%
REC	*sam*, 'put'	10	0.61%

Continued on next page

Table A.2 – *Continued from previous page*

Dative function	Predicate	Frequency	% of Dative function
REC	*huxzar*, 'be returned'	7	0.43%
REC	*masar*, 'deliver'	7	0.43%
REC	*xilek*, 'divide, hand out'	7	0.43%
REC	*hugash*, 'be served'	6	0.37%
REC	*taram*, 'donate'	6	0.37%
REC	*hikna*, 'instill'	5	0.30%
REC	*huanak*, 'be granted'	5	0.30%
REC	*huavar*, 'be transfered'	5	0.30%
REC	*shulam*, 'be paid'	5	0.30%
REC	*hikciv*, 'allocate'	3	0.18%
REC	*himci*, 'invent'	3	0.18%
REC	*hinpik*, 'issue'	3	0.18%
REC	*hukcav*, 'be allocated'	3	0.18%
REC	*hukna*, 'be instilled'	3	0.18%
REC	*husaf*, 'be added'	3	0.18%
REC	*xulak*, 'be divided, handed out'	3	0.18%
REC	*heenik*, 'grant'	2	0.12%
REC	*hilva*, 'lend'	2	0.12%
REC	*hixnis*, 'insert'	2	0.12%
REC	*hunpak*, 'be issued'	2	0.12%
REC	*nishar*, 'be left'	2	0.12%
REC	*supak*, 'be supplied'	2	0.12%
REC	*xayav*, 'owe'	2	0.12%
REC	*braxot*, 'congratulations'	1	0.06%
REC	*hefic*, 'distribute'	1	0.06%
REC	*hefik*, 'produce'	1	0.06%
REC	*heshiv*, 'return, answer, reply'	1	0.06%
REC	*hiciv*, 'place'	1	0.06%
REC	*higia*, 'deserve'	1	0.06%
REC	*hizrik*, 'inject'	1	0.06%
REC	*hoci*, 'take out'	1	0.06%
REC	*taman*, 'hide'	1	0.06%

B

Multiple Correspondence Analysis: numerical summary

```
> dimdesc(res.mca)
$`Dim 1`
$`Dim 1`$quali
                          R2       p.value
DAT.FUNC         0.681176942  0.000000e+00
NO.D             0.450909844  0.000000e+00
NO.OF.ARGUMENTS 0.664632560  0.000000e+00
O.DEF            0.354169649  0.000000e+00
O.INDIV          0.354212240  0.000000e+00
P.LEX.CAT        0.811319987  0.000000e+00
P.TYPE           0.685560552  0.000000e+00
PRED.FIRST       0.729449133  0.000000e+00
S.AGE.INDIV      0.748186807  0.000000e+00
V.BIN            0.773231272  0.000000e+00
DAT.PRSN         0.078629562 4.637403e-173
MODE             0.036162821  1.288879e-79
AFFIRM           0.004550639  2.960016e-11
P.VOICE          0.003627648  2.942905e-09
ELLIPSIS         0.001564069  9.824994e-05
```

$‘Dim 1‘$category

	Estimate	p.value
V.BIN_IRR	0.98278160	0.000000e+00
S.irr	1.16003484	0.000000e+00
predFirst	0.69289188	0.000000e+00
P.TYPE_IRR	1.22809686	0.000000e+00
O.def.irr	0.56353889	0.000000e+00
one	1.04124803	0.000000e+00
NO.D_IRR	1.25446067	0.000000e+00
EXP	0.69462592	0.000000e+00
EVAL	1.01495279	0.000000e+00
DM	0.83799007	0.000000e+00
P.TYPE_PROPERTY	0.55789196	8.352198e-245
O.irr	0.66529224	1.159751e-217
frst	0.23749094	7.388190e-169
P.TYPE_SD	0.44528198	4.580598e-157
V.BIN_NIFAL	0.34591911	5.082963e-114
P.TYPE_VALUE	0.41027185	1.032523e-102
P.LEX.CAT_N	0.72846306	3.054817e-84
irrealis	0.13768083	1.288879e-79
V.BIN_HITPAEL	0.31268823	6.638144e-27
P.LEX.CAT_D.M	0.71662759	1.492345e-16
negative	0.08339664	2.960016e-11
P.LEX.CAT_P	0.71517709	2.785490e-10
pass/mid	0.12072082	2.942905e-09
P.LEX.CAT_ADV	0.21877618	9.027713e-08
P.LEX.CAT_INTERROGATIVE	0.39388461	1.007188e-07
ellip.yes	0.07237028	9.824994e-05
V.BIN_PUAL	-0.13550130	6.788613e-03
NO.D_V	-0.04818848	4.155820e-03
O.high	-0.19329833	1.933982e-04
P.LEX.CAT_COMP.V	-0.66327774	1.561202e-04
ellip.no	-0.07237028	9.824994e-05
P.TYPE_SA	-0.23827583	7.896080e-05
ETH	-0.20591999	7.956395e-06
P.TYPE_SB	-0.09943628	3.505607e-07
active	-0.12072082	2.942905e-09
O.mid-low	-0.17544383	2.288703e-09
affirmative	-0.08339664	2.960016e-11
O.mid-high	-0.18286399	6.703168e-14
scnd	-0.09032158	2.565094e-21

```
NO.D_P                  -0.24156603  2.173574e-23
two                     -0.11183998  5.833445e-35
thrd                    -0.14716936  1.592580e-59
realis                  -0.13768083  1.288879e-79
AFF                     -0.55034432  2.261549e-94
P.LEX.CAT_INTRANS.V     -0.70968668 1.132288e-118
POSSDAT                 -0.55046856 1.191647e-126
NO.D_CL                 -0.29761118 2.765395e-133
O.def.yes               -0.28752872 1.116309e-142
O.def.no                -0.27601017 1.412301e-155
HEP                     -0.29707835 3.156061e-158
ADD                     -0.31529671 4.560717e-222
P.LEX.CAT_TRANS.V       -1.43114356 6.561886e-244
V.BIN_PIEL              -0.52275705 1.328146e-313
V.BIN_KAL               -0.54827995  0.000000e+00
V.BIN_HIFIL             -0.44888934 1.245045e-321
S.mid                   -0.43780392  0.000000e+00
S.low                   -0.28242076  0.000000e+00
S.high                  -0.43981016  0.000000e+00
predNotFirst            -0.69289188  0.000000e+00
P.TYPE_SC               -0.67786599  0.000000e+00
P.TYPE_PB               -0.69860038  0.000000e+00
P.TYPE_PA               -0.92736416  0.000000e+00
three                   -0.92940805  0.000000e+00
NO.D_NP                 -0.67338816  0.000000e+00
REC                     -0.62846084  0.000000e+00


$`Dim 2`
$`Dim 2`$quali
                             R2       p.value
DAT.FUNC          0.482422190  0.000000e+00
MODE              0.188777693  0.000000e+00
NO.D              0.718022346  0.000000e+00
O.DEF             0.415886642  0.000000e+00
O.INDIV           0.415498053  0.000000e+00
P.TYPE            0.573390670  0.000000e+00
S.AGE.INDIV       0.253678916  0.000000e+00
V.BIN             0.128231609 6.694675e-283
P.LEX.CAT         0.117452568 5.317893e-256
NO.OF.ARGUMENTS   0.110282407 1.263077e-246
```

```
DAT.PRSN        0.084458947 2.041224e-186
PRED.FIRST      0.029631551  2.325301e-65
AFFIRM          0.027408142  1.585032e-60
ELLIPSIS        0.002997156  6.907301e-08

$`Dim 2`$category
                       Estimate       p.value
O.def.yes            0.31155658  0.000000e+00
NO.D_NP              0.39971428  0.000000e+00
NO.D_IRR             0.80075787  0.000000e+00
realis               0.23078102  0.000000e+00
P.TYPE_IRR           0.83299027 5.178253e-269
S.mid                0.35355682 3.345823e-253
one                  0.47175893 1.216497e-229
thrd                 0.18528013 2.937007e-168
P.TYPE_PA            0.23337167 2.910029e-136
O.def.no             0.15804379 1.177079e-106
REC                  0.27436841 5.213566e-105
predFirst            0.10245369  2.325301e-65
V.BIN_IRR            0.28609536  5.479213e-61
negative             0.15015319  1.585032e-60
NO.D_V               0.13045624  8.705716e-49
POSSDAT              0.27455167  4.000868e-38
O.mid-high           0.21189099  2.329949e-35
AFF                  0.24050155  2.529491e-22
DM                   0.16612253  4.842442e-19
P.LEX.CAT_N          0.51218726  3.763412e-18
P.TYPE_PROPERTY      0.09218768  3.038311e-11
ellip.yes            0.07349696  6.907301e-08
O.high               0.19157661  1.216151e-07
P.TYPE_VALUE         0.07789804  1.323856e-06
V.BIN_PIEL           0.07910764  3.584116e-05
EVAL                 0.06103740  2.031218e-04
O.mid-low            0.07456255  2.717218e-04
V.BIN_HUFAL          0.16436501  2.748496e-04
P.LEX.CAT_P          0.59816775  8.721272e-04
P.LEX.CAT_D.M        0.45477703  9.416551e-04
frst                -0.01326457  3.095173e-02
V.BIN_HITPAEL       -0.10744980  1.008762e-02
ellip.no            -0.07349696  6.907301e-08
P.TYPE_SC           -0.07500686  8.573691e-11
```

```
P.TYPE_SB           -0.11379215  9.220515e-12
P.LEX.CAT_ADJ       -0.32746214  8.576862e-12
P.LEX.CAT_TRANS.V   -0.57011967  7.710633e-18
V.BIN_HIFIL         -0.15999281  7.005710e-23
P.LEX.CAT_INTRANS.V -0.66536731  2.251917e-43
HEP                 -0.16436993  5.270911e-58
affirmative         -0.15015319  1.585032e-60
predNotFirst        -0.10245369  2.325301e-65
V.BIN_NIFAL         -0.38046577  3.562355e-68
two                 -0.20063239  7.681969e-76
NO.D_P              -0.24228370  1.545667e-79
EXP                 -0.21071766  7.641231e-94
scnd                -0.17201557 4.050116e-131
three               -0.27112654 1.976523e-222
P.TYPE_SD           -0.48100725 1.425223e-244
NO.D_A              -0.67835269 6.652929e-276
O.irr               -0.54679398 2.049830e-295
S.low               -0.36659009  0.000000e+00
P.TYPE_PB           -0.52343700  0.000000e+00
O.def.irr           -0.46960038  0.000000e+00
NO.D_CL             -0.41029200  0.000000e+00
irrealis            -0.23078102  0.000000e+00
ADD                 -0.59114641  0.000000e+00


$`Dim 3`
$`Dim 3`$quali
                           R2       p.value
DAT.FUNC        0.255857146  0.000000e+00
NO.D            0.344404709  0.000000e+00
NO.OF.ARGUMENTS 0.341545654  0.000000e+00
P.LEX.CAT       0.709888342  0.000000e+00
P.TYPE          0.463872904  0.000000e+00
P.VOICE         0.263257559  0.000000e+00
V.BIN           0.510097573  0.000000e+00
DAT.PRSN        0.127946130 7.991456e-289
MODE            0.078792278 5.454314e-175
AFFIRM          0.023949565  5.009959e-53
S.AGE.INDIV     0.022424654  2.240542e-47
O.DEF           0.003275552  1.246522e-07
O.INDIV         0.003033139  6.375225e-06
```

```
PRED.FIRST      0.001332303  3.249818e-04

$‘Dim 3‘$category
                           Estimate       p.value
pass/mid                 0.73299851  0.000000e+00
P.TYPE_SD                0.64188875  0.000000e+00
P.LEX.CAT_INTRANS.V      1.42686216  0.000000e+00
two                      0.51036425  0.000000e+00
NO.D_A                   1.21974669  0.000000e+00
V.BIN_NIFAL              0.55239914 1.267440e-253
V.BIN_HUFAL              0.97562796 4.000784e-185
realis                   0.14485291 5.454314e-175
EXP                      0.29478483 3.168333e-133
frst                     0.13836306 3.798690e-121
P.LEX.CAT_ADJ            0.59521645 1.463403e-107
V.BIN_PUAL               0.95982869  1.245307e-73
thrd                     0.10713200  2.519093e-65
P.TYPE_PA                0.16960758  8.569915e-63
S.mid                    0.18420866  9.837706e-60
P.TYPE_SB                0.29599084  5.642200e-59
negative                 0.13636541  5.009959e-53
NO.D_NP                  0.11285342  6.979652e-39
P.LEX.CAT_TRANS.V        0.46315842  5.705999e-36
P.LEX.CAT_ADV            0.37299684  7.760846e-25
P.TYPE_SA                0.54045877  7.876996e-22
three                    0.05582281  9.048165e-16
ETH                      0.39581810  3.391988e-15
P.LEX.CAT_COMP.V         0.98406510  2.253056e-10
REC                      0.08635074  2.734488e-09
POSSDAT                  0.13366069  6.164213e-08
O.def.yes                0.05248296  9.962805e-08
P.TYPE_PROPERTY          0.07943276  1.447949e-07
P.TYPE_VALUE             0.07250004  3.575188e-05
predFirst                0.02110632  3.249818e-04
P.TYPE_SC                0.04211591  8.120825e-04
O.mid-high               0.05621460  9.200907e-03
V.BIN_HITPAEL            0.06727414  2.697728e-02
predNotFirst            -0.02110632  3.249818e-04
NO.D_V                  -0.04785132  2.584017e-04
S.low                   -0.03399622  2.166569e-06
O.def.no                -0.05091328  1.766961e-08
```

S.high	-0.06247859	2.652238e-12
P.LEX.CAT_INTERROGATIVE	-0.54245088	1.136963e-16
S.irr	-0.08773385	8.690622e-19
P.LEX.CAT_P	-0.99329759	4.103985e-23
NO.D_P	-0.26157623	1.886724e-43
P.LEX.CAT_D.M	-1.08563365	3.517125e-45
affirmative	-0.13636541	5.009959e-53
P.TYPE_PB	-0.21240826	2.600008e-102
ADD	-0.25154214	5.398973e-123
NO.D_CL	-0.22402257	6.965870e-125
irrealis	-0.14485291	5.454314e-175
DM	-0.65929016	2.046866e-194
P.LEX.CAT_N	-1.22091685	7.449107e-284
scnd	-0.24549506	7.416520e-284
V.BIN_PIEL	-0.55821529	0.000000e+00
V.BIN_KAL	-0.62120863	0.000000e+00
V.BIN_IRR	-0.71140747	0.000000e+00
V.BIN_HIFIL	-0.66429853	0.000000e+00
active	-0.73299851	0.000000e+00
P.TYPE_IRR	-1.62958639	0.000000e+00
one	-0.56618707	0.000000e+00
NO.D_IRR	-0.79914999	0.000000e+00

C

Hierarchical Clustering on Principal Components: description of clusters by individuals and variables

C.1 Description of the clusters according to individuals: central and unique exemplars

```
> res.hcpc.default$desc.ind
$para
cluster: 1
       189        191       2233       4308       5230
0.07030954 0.07030954 0.07030954 0.07030954 0.07030954
---------------------------------------------------------------------------
cluster: 2
      130       686      1319      1463      4077
0.1044625 0.1044625 0.1044625 0.1044625 0.1044625
---------------------------------------------------------------------------
cluster: 3
      624     10681     12931       375      1078
0.1639668 0.1720700 0.1860599 0.1912814 0.1912814
---------------------------------------------------------------------------
```

```
cluster: 4
     4257      1971      2138     12585     15686
0.1174934 0.1230346 0.1230346 0.1230346 0.1230346
------------------------------------------------------------------------
cluster: 5
     4941       356      2468     16281     16664
0.3122393 0.4203189 0.4203189 0.4203189 0.4203189

$dist
cluster: 1
   12270      109    16060    16311     8362
1.741606 1.741606 1.717634 1.717634 1.689767
------------------------------------------------------------------------
cluster: 2
    8948     8191       45       15       30
1.341793 1.325959 1.291593 1.291593 1.291593
------------------------------------------------------------------------
cluster: 3
   10730    11102    11211     9253    11419
2.712026 2.644094 2.644094 2.557724 2.557724
------------------------------------------------------------------------
cluster: 4
    5937     7124     8602    14658      975
1.905856 1.905856 1.905856 1.896368 1.895277
------------------------------------------------------------------------
cluster: 5
    623     684     757     984    1155
3.65234 3.65234 3.65234 3.65234 3.65234
```

C.2 Description of clusters according to variables

```
> res.hcpc.default$desc.var
$test.chi2
                      p.value df
DAT.FUNC         0.000000e+00 32
NO.D             0.000000e+00 20
NO.OF.ARGUMENTS  0.000000e+00  8
O.DEF            0.000000e+00  8
O.INDIV          0.000000e+00 16
P.LEX.CAT        0.000000e+00 32
P.TYPE           0.000000e+00 32
P.VOICE          0.000000e+00  4
PRED.FIRST       0.000000e+00  4
S.AGE.INDIV      0.000000e+00 12
V.BIN            0.000000e+00 28
DAT.PRSN        3.734491e-273  8
MODE            1.886417e-269  4
ELLIPSIS         1.582315e-52  4
AFFIRM           1.649393e-40  4

$category
$category$‘1‘
                                  Cla/Mod     Mod/Cla      Global      p.value     v.test
P.TYPE=P.TYPE_PA              69.85743381 71.70731707 30.38993192 0.000000e+00        Inf
P.LEX.CAT=P.LEX.CAT_TRANS.V   41.00014411 99.12891986 71.58035898 0.000000e+00        Inf
O.INDIV=O.mid-low             99.72421401 62.99651568 18.70229008 0.000000e+00        Inf
O.INDIV=O.mid-high           100.00000000 28.53658537  8.44852486 0.000000e+00        Inf
O.DEF=O.def.yes              100.00000000 23.37979094  6.92180730 0.000000e+00        Inf
O.DEF=O.def.no                99.61464355 72.05574913 21.41530844 0.000000e+00        Inf
NO.OF.ARGUMENTS=three         45.03903709 96.48083624 63.42067258 0.000000e+00        Inf
NO.D=NO.D_NP                  59.34295878 98.81533101 49.29853518 0.000000e+00        Inf
```

DAT.FUNC=REC	89.09866017	50.97560976	16.93831236	0.000000e+00	Inf
PRED.FIRST=predNotFirst	36.52129559	99.79094077	80.89539922	3.991799e-306	37.399429
MODE=realis	52.93657009	47.10801394	26.34619352	1.710959e-188	29.282930
V.BIN=V.BIN_KAL	43.50011970	63.31010453	43.08850836	1.048012e-149	26.059572
S.AGE.INDIV=S.mid	71.35862913	20.31358885	8.42789354	1.813164e-147	25.861357
DAT.PRSN=thrd	43.07592050	46.06271777	31.65875799	2.262103e-84	19.462987
S.AGE.INDIV=S.high	46.40338505	34.39024390	21.94140706	6.708147e-78	18.683791
DAT.FUNC=POSSDAT	81.69642857	6.37630662	2.31070765	5.351903e-60	16.337362
DAT.FUNC=AFF	85.53459119	4.73867596	1.64018981	1.954324e-49	14.780595
P.VOICE=active	30.39610252	100.00000000	97.40045389	9.341166e-40	13.195276
O.INDIV=O.high	100.00000000	2.40418118	0.71178048	1.873409e-37	12.789639
O.INDIV=O.low	93.47826087	1.49825784	0.47452032	8.441168e-20	9.107359
ELLIPSIS=ellip.no	30.21299255	98.85017422	96.86403961	2.203232e-15	7.929335
V.BIN=V.BIN_PIEL	39.22462942	11.98606272	9.04683309	1.667222e-10	6.389231
AFFIRM=negative	35.53956835	8.60627178	7.16938312	4.626396e-04	3.501504
P.LEX.CAT=P.LEX.CAT_D.M	0.00000000	0.00000000	0.09284093	4.237847e-02	-2.029784
P.TYPE=P.TYPE_SA	10.71428571	0.10452962	0.28883846	2.153101e-02	-2.298541
P.LEX.CAT=P.LEX.CAT_INTERROGATIVE	0.00000000	0.00000000	0.13410357	1.038716e-02	-2.562667
P.TYPE=P.TYPE_SC	25.22432702	8.81533101	10.34660615	1.160672e-03	-3.248373
AFFIRM=affirmative	29.14768308	91.39372822	92.83061688	4.626396e-04	-3.501504
V.BIN=V.BIN_PUAL	0.00000000	0.00000000	0.25789148	1.523126e-04	-3.787268
V.BIN=V.BIN_HITPAEL	4.30107527	0.13937282	0.95935630	6.472070e-10	-6.178482
P.LEX.CAT=P.LEX.CAT_ADV	0.00000000	0.00000000	0.76335878	4.639368e-12	-6.916195
DAT.PRSN=scnd	24.08963585	20.97560976	25.77883227	1.229230e-12	-7.102046
V.BIN=V.BIN_HUFAL	0.00000000	0.00000000	0.80462142	1.124238e-12	-7.114372
ELLIPSIS=ellip.yes	10.85526316	1.14982578	3.13596039	2.203232e-15	-7.929335
P.LEX.CAT=P.LEX.CAT_N	0.00000000	0.00000000	1.20693212	1.077948e-18	-8.826715
NO.D=NO.D_A	0.00000000	0.00000000	1.34103569	1.047049e-20	-9.331174
P.TYPE=P.TYPE_IRR	0.00000000	0.00000000	1.59892717	1.381236e-24	-10.235050
S.AGE.INDIV=S.low	24.91369390	45.26132404	53.78584692	1.059215e-27	-10.907683

```
DAT.PRSN=frst                      22.92777508  32.96167247 42.56240974  7.588480e-36 -12.498697
P.VOICE=pass/mid                    0.00000000   0.00000000  2.59954611  9.341166e-40 -13.195276
NO.D=NO.D_P                         1.32450331   0.13937282  3.11532907  1.611402e-40 -13.327052
DAT.FUNC=DM                         0.00000000   0.00000000  3.12564473  8.298326e-48 -14.525950
P.TYPE=P.TYPE_SB                    0.00000000   0.00000000  3.45574582  6.920588e-53 -15.306483
DAT.FUNC=EVAL                       0.00000000   0.00000000  4.22942026  7.144009e-65 -17.008165
NO.D=NO.D_IRR                       0.00000000   0.00000000  4.63173097  3.756109e-71 -17.835362
P.TYPE=P.TYPE_VALUE                 0.00000000   0.00000000  4.67299360  8.489934e-72 -17.918289
NO.OF.ARGUMENTS=one                 0.00000000   0.00000000  4.68330926  5.853194e-72 -17.938968
P.TYPE=P.TYPE_PROPERTY              0.36101083   0.06968641  5.71487518  1.050265e-83 -19.384150
P.TYPE=P.TYPE_SD                    0.00000000   0.00000000  5.56014029  9.140234e-86 -19.626735
V.BIN=V.BIN_NIFAL                   0.00000000   0.00000000  5.57045595  6.274863e-86 -19.645841
DAT.FUNC=ADD                       16.56497865  18.91986063 33.81473076  1.447280e-95 -20.742017
NO.D=NO.D_V                         2.69058296   1.04529617 11.50195998 7.879063e-134 -24.619148
P.TYPE=P.TYPE_PB                   15.05025808  19.30313589 37.97194141 6.136796e-142 -25.365018
P.LEX.CAT=P.LEX.CAT_INTRANS.V       1.52979066   0.66202091 12.81204869 3.599570e-170 -27.806700
MODE=irrealis                      21.26050420  52.89198606 73.65380648 1.710959e-188 -29.282930
P.LEX.CAT=P.LEX.CAT_ADJ             0.46403712   0.20905923 13.33814731 1.975598e-201 -30.282051
V.BIN=V.BIN_IRR                     0.39630119   0.20905923 15.61790798 1.066581e-241 -33.200280
S.AGE.INDIV=S.irr                   0.06510417   0.03484321 15.84485249 1.572339e-257 -34.279749
DAT.FUNC=EXP                        1.09589041   0.69686411 18.82607799 3.647791e-274 -35.378567
PRED.FIRST=predFirst                0.32397408   0.20905923 19.10460078 3.991799e-306 -37.399429
O.INDIV=O.irr                       1.88570606   4.56445993 71.66288426  0.000000e+00        -Inf
O.DEF=O.def.irr                     1.88570606   4.56445993 71.66288426  0.000000e+00        -Inf
NO.OF.ARGUMENTS=two                 3.26649418   3.51916376 31.89601816  0.000000e+00        -Inf
NO.D=NO.D_CL                        0.00000000   0.00000000 30.11140912  0.000000e+00        -Inf

$category$‘2‘
                                    Cla/Mod      Mod/Cla      Global      p.value     v.test
PRED.FIRST=predNotFirst           52.0785514  98.95808093 80.89539922  0.000000e+00        Inf
```

P.TYPE=P.TYPE_PB	78.5656072	70.07511510	37.97194141	0.000000e+00	Inf
P.LEX.CAT=P.LEX.CAT_TRANS.V	58.7404525	98.76423552	71.58035898	0.000000e+00	Inf
O.INDIV=O.irr	59.3061753	99.83038527	71.66288426	0.000000e+00	Inf
O.DEF=O.def.irr	59.3061753	99.83038527	71.66288426	0.000000e+00	Inf
NO.D=NO.D_CL	80.1644399	56.69978192	30.11140912	0.000000e+00	Inf
DAT.FUNC=ADD	80.9334960	64.28398352	33.81473076	0.000000e+00	Inf
S.AGE.INDIV=S.low	57.3647871	72.47395202	53.78584692	3.291615e-227	32.180656
MODE=irrealis	51.5126050	89.12042646	73.65380648	4.617492e-210	30.930711
NO.OF.ARGUMENTS=three	52.7813923	78.62854374	63.42067258	2.706845e-163	27.232115
DAT.PRSN=scnd	62.9051621	38.09062273	25.77883227	3.519300e-125	23.797804
V.BIN=V.BIN_HIFIL	63.2635983	36.63678217	24.65442542	3.942339e-122	23.501445
DAT.FUNC=HEP	61.5853659	26.92028108	18.60944914	1.512998e-72	18.013997
P.VOICE=active	43.7089600	100.00000000	97.40045389	1.666573e-62	16.685659
P.TYPE=P.TYPE_SC	64.6061815	15.70147807	10.34660615	9.945290e-50	14.826023
V.BIN=V.BIN_KAL	50.7062485	51.32057184	43.08850836	4.433733e-45	14.089068
AFFIRM=affirmative	44.3382598	96.68039738	92.83061688	9.809368e-40	13.191590
NO.D=NO.D_V	59.6412556	16.11339956	11.50195998	4.368416e-34	12.172256
NO.D=NO.D_P	74.8344371	5.47613278	3.11532907	8.117992e-31	11.541830
ELLIPSIS=ellip.yes	73.3552632	5.40344076	3.13596039	2.742467e-28	11.029875
S.AGE.INDIV=S.high	50.8227551	26.19336079	21.94140706	4.246380e-18	8.671989
V.BIN=V.BIN_PIEL	55.9863170	11.89726193	9.04683309	6.049039e-17	8.364266
P.LEX.CAT=P.LEX.CAT_D.M	0.0000000	0.00000000	0.09284093	6.774344e-03	-2.707738
P.LEX.CAT=P.LEX.CAT_INTERROGATIVE	0.0000000	0.00000000	0.13410357	7.344176e-04	-3.376397
P.TYPE=P.TYPE_SA	3.5714286	0.02423068	0.28883846	3.997334e-06	-4.611521
P.TYPE=P.TYPE_SB	29.8507463	2.42306760	3.45574582	1.111834e-06	-4.870736
V.BIN=V.BIN_PUAL	0.0000000	0.00000000	0.25789148	9.288990e-07	-4.906132
DAT.FUNC=ETH	6.1224490	0.07269203	0.50546730	1.305659e-08	-5.685323
O.INDIV=O.low	4.3478261	0.04846135	0.47452032	4.976375e-09	-5.847960
O.INDIV=O.high	0.0000000	0.00000000	0.71178048	1.998638e-17	-8.493872
P.LEX.CAT=P.LEX.CAT_ADV	0.0000000	0.00000000	0.76335878	1.214549e-18	-8.813355

V.BIN=V.BIN_HUFAL	0.0000000	0.00000000	0.80462142	1.290475e-19	-9.061179
V.BIN=V.BIN_HITPAEL	2.1505376	0.04846135	0.95935630	7.326894e-20	-9.122710
DAT.PRSN=frst	36.6214251	36.61255149	42.56240974	1.519604e-24	-10.225805
DAT.FUNC=AFF	5.0314465	0.19384541	1.64018981	2.083448e-27	-10.846000
ELLIPSIS=ellip.no	41.5761448	94.59655924	96.86403961	2.742467e-28	-11.029875
P.LEX.CAT=P.LEX.CAT_N	0.0000000	0.00000000	1.20693212	3.881030e-29	-11.204370
DAT.PRSN=thrd	34.0175953	25.29682578	31.65875799	1.811807e-31	-11.670114
NO.D=NO.D_A	0.0000000	0.00000000	1.34103569	2.531841e-32	-11.836367
P.TYPE=P.TYPE_IRR	0.0000000	0.00000000	1.59892717	1.824172e-38	-12.969415
P.TYPE=P.TYPE_VALUE	14.7902870	1.62345529	4.67299360	1.048855e-38	-13.011769
AFFIRM=negative	19.7122302	3.31960262	7.16938312	9.809368e-40	-13.191590
DAT.FUNC=POSSDAT	4.0178571	0.21807608	2.31070765	5.582091e-41	-13.405928
P.VOICE=pass/mid	0.0000000	0.00000000	2.59954611	1.666573e-62	-16.685659
NO.OF.ARGUMENTS=two	28.5252264	21.37145626	31.89601816	7.840010e-84	-19.399188
DAT.FUNC=EVAL	0.0000000	0.00000000	4.22942026	2.194282e-102	-21.484067
NO.D=NO.D_IRR	0.0000000	0.00000000	4.63173097	2.266014e-112	-22.526831
NO.OF.ARGUMENTS=one	0.0000000	0.00000000	4.68330926	1.177201e-113	-22.657482
P.TYPE=P.TYPE_SD	0.0000000	0.00000000	5.56014029	1.229043e-135	-24.787295
S.AGE.INDIV=S.mid	5.7527540	1.13884177	8.42789354	1.107733e-135	-24.791481
V.BIN=V.BIN_NIFAL	0.0000000	0.00000000	5.57045595	6.750005e-136	-24.811421
P.TYPE=P.TYPE_PROPERTY	0.0000000	0.00000000	5.71487518	1.518891e-139	-25.147137
O.DEF=O.def.yes	0.0000000	0.00000000	6.92180730	2.474972e-170	-27.820151
MODE=realis	17.5802662	10.87957354	26.34619352	4.617492e-210	-30.930711
O.INDIV=O.mid-high	0.0000000	0.00000000	8.44852486	4.465323e-210	-30.931793
P.LEX.CAT=P.LEX.CAT_INTRANS.V	3.8647343	1.16307245	12.81204869	4.053930e-243	-33.298537
DAT.FUNC=EXP	7.2328767	3.19844924	18.82607799	1.875622e-299	-36.986671
DAT.FUNC=REC	4.3848965	1.74460867	16.93831236	1.828043e-321	-38.331549
V.BIN=V.BIN_IRR	0.2642008	0.09692270	15.61790798	0.000000e+00	-Inf
S.AGE.INDIV=S.irr	0.5208333	0.19384541	15.84485249	0.000000e+00	-Inf
PRED.FIRST=predFirst	2.3218143	1.04191907	19.10460078	0.000000e+00	-Inf

```
P.TYPE=P.TYPE_PA                   14.2226748  10.15265326 30.38993192  0.000000e+00        -Inf
P.LEX.CAT=P.LEX.CAT_ADJ             0.2320186   0.07269203 13.33814731  0.000000e+00        -Inf
O.INDIV=O.mid-low                   0.2757860   0.12115338 18.70229008  0.000000e+00        -Inf
O.DEF=O.def.no                      0.3371869   0.16961473 21.41530844  0.000000e+00        -Inf
NO.D=NO.D_NP                       18.7486922  21.71068573 49.29853518  0.000000e+00        -Inf

$category$`3`
                                   Cla/Mod      Mod/Cla      Global       p.value     v.test
V.BIN=V.BIN_NIFAL                98.14814815  48.18181818  5.5704560  0.000000e+00        Inf
P.LEX.CAT=P.LEX.CAT_INTRANS.V    88.16425121  99.54545455 12.8120487  0.000000e+00        Inf
NO.OF.ARGUMENTS=two              30.14230272  84.72727273 31.8960182  0.000000e+00        Inf
DAT.FUNC=EXP                     39.01369863  64.72727273 18.8260780 8.548797e-290  36.380855
P.VOICE=pass/mid                100.00000000  22.90909091  2.5995461 4.978060e-251  33.840671
O.INDIV=O.irr                    15.81977832  99.90909091 71.6628843 5.608365e-169  27.707898
O.DEF=O.def.irr                  15.81977832  99.90909091 71.6628843 5.608365e-169  27.707898
P.TYPE=P.TYPE_SD                 56.40074212  27.63636364  5.5601403 2.587632e-156  26.636181
NO.D=NO.D_A                     100.00000000  11.81818182  1.3410357 1.159679e-126  23.940530
V.BIN=V.BIN_HUFAL               100.00000000   7.09090909  0.8046214  1.591465e-75  18.389601
PRED.FIRST=predFirst             24.29805616  40.90909091 19.1046008  5.778071e-72  17.939686
V.BIN=V.BIN_HITPAEL              86.02150538   7.27272727  0.9593563  1.125847e-62  16.709066
P.TYPE=P.TYPE_SB                 44.47761194  13.54545455  3.4557458  1.843750e-55  15.687428
NO.D=NO.D_NP                     14.81481481  64.36363636 49.2985352  1.430071e-26  10.668440
V.BIN=V.BIN_PUAL                100.00000000   2.27272727  0.2578915  1.846545e-24  10.206909
DAT.PRSN=frst                    14.97818711  56.18181818 42.5624097  5.570760e-22   9.637160
MODE=realis                      16.13155834  37.45454545 26.3461935  7.916862e-18   8.600791
ELLIPSIS=ellip.no                11.71458999 100.00000000 96.8640396  6.735284e-17   8.351585
DAT.FUNC=ETH                     59.18367347   2.63636364  0.5054673  8.923162e-16   8.040830
AFFIRM=negative                  21.29496403  13.45454545  7.1693831  4.373123e-15   7.843743
P.TYPE=P.TYPE_SA                 71.42857143   1.81818182  0.2888385  1.423014e-13   7.394163
S.AGE.INDIV=S.irr                17.05729167  23.81818182 15.8448525  2.451398e-13   7.321541
```

S.AGE.INDIV=S.low	13.27196011	62.90909091	53.7858469	8.909291e-11	6.484388
P.TYPE=P.TYPE_PA	13.81534284	37.00000000	30.3899319	6.230062e-07	4.983946
DAT.PRSN=thrd	12.51221896	34.90909091	31.6587580	1.449738e-02	2.444697
O.INDIV=O.low	2.17391304	0.09090909	0.4745203	3.065865e-02	-2.161475
DAT.FUNC=AFF	5.66037736	0.81818182	1.6401898	1.508905e-02	-2.430235
NO.OF.ARGUMENTS=one	7.48898678	3.09090909	4.6833093	5.542882e-03	-2.773664
NO.D=NO.D_IRR	6.45879733	2.63636364	4.6317310	3.706181e-04	-3.560168
O.INDIV=O.high	0.00000000	0.00000000	0.7117805	2.383695e-04	-3.674442
P.LEX.CAT=P.LEX.CAT_ADV	0.00000000	0.00000000	0.7633588	1.299154e-04	-3.826619
NO.D=NO.D_P	4.96688742	1.36363636	3.1153291	1.021764e-04	-3.885365
P.LEX.CAT=P.LEX.CAT_N	0.00000000	0.00000000	1.2069321	6.930353e-07	-4.963305
DAT.FUNC=DM	3.30033003	0.90909091	3.1256447	3.170617e-07	-5.113012
P.TYPE=P.TYPE_IRR	0.64516129	0.09090909	1.5989272	1.477305e-07	-5.255366
DAT.FUNC=HEP	7.64966741	12.54545455	18.6094491	1.097661e-08	-5.714904
V.BIN=V.BIN_PIEL	4.67502851	3.72727273	9.0468331	5.670305e-13	-7.208184
DAT.FUNC=REC	6.33373934	9.45454545	16.9383124	7.902054e-14	-7.471941
AFFIRM=affirmative	10.57895322	86.54545455	92.8306169	4.373123e-15	-7.843743
P.TYPE=P.TYPE_SC	4.38683948	4.00000000	10.3466061	3.876496e-16	-8.142359
ELLIPSIS=ellip.yes	0.00000000	0.00000000	3.1359604	6.735284e-17	-8.351585
MODE=irrealis	9.63585434	62.54545455	73.6538065	7.916862e-18	-8.600791
DAT.FUNC=EVAL	0.24390244	0.09090909	4.2294203	6.528265e-21	-9.381109
P.TYPE=P.TYPE_VALUE	0.22075055	0.09090909	4.6729936	3.148442e-23	-9.927891
V.BIN=V.BIN_HIFIL	5.85774059	12.72727273	24.6544254	3.762808e-25	-10.360173
NO.D=NO.D_CL	6.37204522	16.90909091	30.1114091	2.247677e-26	-10.626337
NO.D=NO.D_V	2.86995516	2.90909091	11.5019600	1.017804e-27	-10.911309
O.DEF=O.def.yes	0.00000000	0.00000000	6.9218073	3.512548e-37	-12.740694
O.INDIV=O.mid-high	0.00000000	0.00000000	8.4485249	1.305568e-45	-14.175154
DAT.PRSN=scnd	3.92156863	8.90909091	25.7788323	3.086301e-50	-14.904387
P.LEX.CAT=P.LEX.CAT_ADJ	0.07733952	0.09090909	13.3381473	2.306662e-71	-17.862594
S.AGE.INDIV=S.high	1.97461213	3.81818182	21.9414071	1.433772e-71	-17.889111

PRED.FIRST=predNotFirst	8.28870186	59.09090909	80.8953992	5.778071e-72	-17.939686
V.BIN=V.BIN_KAL	4.90782859	18.63636364	43.0885084	1.052311e-73	-18.160928
V.BIN=V.BIN_IRR	0.06605020	0.09090909	15.6179080	6.046000e-85	-19.530487
P.TYPE=P.TYPE_PB	3.17848411	10.63636364	37.9719414	8.886365e-103	-21.526010
O.INDIV=O.mid-low	0.00000000	0.00000000	18.7022901	1.886480e-106	-21.914575
DAT.FUNC=ADD	2.04392923	6.09090909	33.8147308	2.476426e-119	-23.226187
O.DEF=O.def.no	0.04816956	0.09090909	21.4153084	2.169428e-121	-23.428903
P.VOICE=active	8.98114806	77.09090909	97.4004539	4.978060e-251	-33.840671
NO.OF.ARGUMENTS=three	2.17957059	12.18181818	63.4206726	4.294153e-309	-37.581605
P.LEX.CAT=P.LEX.CAT_TRANS.V	0.05764519	0.36363636	71.5803590	0.000000e+00	-Inf

$category$‘4‘

	Cla/Mod	Mod/Cla	Global	p.value	v.test
V.BIN=V.BIN_IRR	72.5231176	96.23137599	15.61790798	0.000000e+00	Inf
S.AGE.INDIV=S.irr	53.5156250	72.04206836	15.84485249	0.000000e+00	Inf
PRED.FIRST=predFirst	48.7580994	79.14110429	19.10460078	0.000000e+00	Inf
P.LEX.CAT=P.LEX.CAT_ADJ	81.6705336	92.55039439	13.33814731	0.000000e+00	Inf
NO.OF.ARGUMENTS=two	36.9016818	100.00000000	31.89601816	0.000000e+00	Inf
DAT.FUNC=EXP	41.3150685	66.08238387	18.82607799	4.764769e-320	38.246450
DAT.FUNC=EVAL	86.8292683	31.20070114	4.22942026	1.321337e-289	36.368893
P.TYPE=P.TYPE_PROPERTY	72.7436823	35.31989483	5.71487518	1.015726e-272	35.284487
P.TYPE=P.TYPE_VALUE	76.3796909	30.32427695	4.67299360	6.646137e-243	33.283701
O.INDIV=O.irr	16.4243558	100.00000000	71.66288426	1.434227e-178	28.492906
O.DEF=O.def.irr	16.4243558	100.00000000	71.66288426	1.434227e-178	28.492906
DAT.PRSN=frst	20.7707222	75.10955302	42.56240974	8.416928e-125	23.761199
NO.D=NO.D_V	33.6322870	32.86590710	11.50195998	3.391922e-97	20.921783
P.TYPE=P.TYPE_SD	33.9517625	16.03856266	5.56014029	9.809762e-45	14.032872
P.LEX.CAT=P.LEX.CAT_ADV	54.0540541	3.50569676	0.76335878	1.012120e-18	8.833763
P.VOICE=active	12.0843042	100.00000000	97.40045389	1.266252e-14	7.709183
ELLIPSIS=ellip.no	12.1299255	99.82471516	96.86403961	1.481171e-14	7.689149

NO.D=NO.D_CL	13.3949983	34.26818580	30.11140912	1.255215e-03	3.226027
AFFIRM=negative	15.3956835	9.37773883	7.16938312	2.934675e-03	2.974499
P.LEX.CAT=P.LEX.CAT_COMP.V	100.0000000	0.17528484	0.02063132	1.384297e-02	2.461313
MODE=irrealis	12.1568627	76.07361963	73.65380648	4.688121e-02	1.987372
MODE=realis	10.6891151	23.92638037	26.34619352	4.688121e-02	-1.987372
V.BIN=V.BIN_PUAL	0.0000000	0.00000000	0.25789148	4.351011e-02	-2.018780
P.TYPE=P.TYPE_SA	0.0000000	0.00000000	0.28883846	2.985170e-02	-2.172053
V.BIN=V.BIN_HITPAEL	4.3010753	0.35056968	0.95935630	1.466129e-02	-2.440639
O.INDIV=O.low	0.0000000	0.00000000	0.47452032	3.105307e-03	-2.957117
AFFIRM=affirmative	11.4901656	90.62226117	92.83061688	2.934675e-03	-2.974499
DAT.FUNC=ETH	0.0000000	0.00000000	0.50546730	2.128648e-03	-3.071670
O.INDIV=O.high	0.0000000	0.00000000	0.71178048	1.711491e-04	-3.758187
V.BIN=V.BIN_HUFAL	0.0000000	0.00000000	0.80462142	5.494845e-05	-4.033516
P.LEX.CAT=P.LEX.CAT_N	0.0000000	0.00000000	1.20693212	3.945498e-07	-5.071569
NO.D=NO.D_A	0.0000000	0.00000000	1.34103569	7.575565e-08	-5.376956
DAT.FUNC=AFF	0.6289308	0.08764242	1.64018981	4.466885e-08	-5.471323
P.TYPE=P.TYPE_IRR	0.0000000	0.00000000	1.59892717	3.149540e-09	-5.923608
DAT.FUNC=POSSDAT	0.0000000	0.00000000	2.31070765	4.635528e-13	-7.235577
ELLIPSIS=ellip.yes	0.6578947	0.17528484	3.13596039	1.481171e-14	-7.689149
P.VOICE=pass/mid	0.0000000	0.00000000	2.59954611	1.266252e-14	-7.709183
P.TYPE=P.TYPE_SC	4.8853440	4.29447853	10.34660615	3.574381e-15	-7.869017
DAT.FUNC=DM	0.0000000	0.00000000	3.12564473	1.745082e-17	-8.509617
NO.D=NO.D_IRR	0.0000000	0.00000000	4.63173097	9.113926e-26	-10.494936
NO.OF.ARGUMENTS=one	0.0000000	0.00000000	4.68330926	4.716305e-26	-10.556972
V.BIN=V.BIN_NIFAL	0.0000000	0.00000000	5.57045595	5.327820e-31	-11.577994
DAT.PRSN=thrd	6.3864451	17.17791411	31.65875799	7.794785e-32	-11.741653
P.LEX.CAT=P.LEX.CAT_INTRANS.V	2.4959742	2.71691499	12.81204869	1.416651e-36	-12.631438
O.DEF=O.def.yes	0.0000000	0.00000000	6.92180730	1.243277e-38	-12.998769
O.INDIV=O.mid-high	0.0000000	0.00000000	8.44852486	2.128761e-47	-14.461256
NO.D=NO.D_NP	6.9052103	28.92199825	49.29853518	5.472189e-50	-14.866084

V.BIN=V.BIN_PIEL	0.0000000	0.00000000	9.04683309	7.076433e-51	-15.002445
DAT.PRSN=scnd	3.5214086	7.71253287	25.77883227	6.398153e-61	-16.466385
DAT.FUNC=HEP	0.9977827	1.57756354	18.60944914	3.260260e-81	-19.086649
DAT.FUNC=REC	0.1218027	0.17528484	16.93831236	1.246898e-94	-20.638172
S.AGE.INDIV=S.high	0.7522332	1.40227870	21.94140706	1.310060e-104	-21.720618
O.INDIV=O.mid-low	0.0000000	0.00000000	18.70229008	1.110626e-110	-22.353732
P.TYPE=P.TYPE_PA	1.8329939	4.73269062	30.38993192	1.756458e-116	-22.942358
V.BIN=V.BIN_HIFIL	0.5020921	1.05170903	24.65442542	1.857741e-128	-24.112292
O.DEF=O.def.no	0.0000000	0.00000000	21.41530844	7.126340e-129	-24.151928
S.AGE.INDIV=S.low	4.2577675	19.45661700	53.78584692	1.632538e-141	-25.326475
P.TYPE=P.TYPE_PB	1.8473241	5.95968449	37.97194141	2.331473e-155	-26.553637
DAT.FUNC=ADD	0.3050641	0.87642419	33.81473076	7.597510e-200	-30.161424
V.BIN=V.BIN_KAL	0.6463969	2.36634531	43.08850836	4.686406e-251	-33.842453
PRED.FIRST=predNotFirst	3.0349401	20.85889571	80.89539922	0.000000e+00	-Inf
P.LEX.CAT=P.LEX.CAT_TRANS.V	0.1729356	1.05170903	71.58035898	0.000000e+00	-Inf
NO.OF.ARGUMENTS=three	0.0000000	0.00000000	63.42067258	0.000000e+00	-Inf

$category$‘5‘

	Cla/Mod	Mod/Cla	Global	p.value	v.test
S.AGE.INDIV=S.irr	28.84114583	97.1491228	15.84485249	0.000000e+00	Inf
PRED.FIRST=predFirst	24.29805616	98.6842105	19.10460078	0.000000e+00	Inf
NO.OF.ARGUMENTS=one	92.51101322	92.1052632	4.68330926	0.000000e+00	Inf
NO.D=NO.D_IRR	93.54120267	92.1052632	4.63173097	0.000000e+00	Inf
V.BIN=V.BIN_IRR	26.75033025	88.8157895	15.61790798	3.879798e-284	36.021319
P.TYPE=P.TYPE_IRR	99.35483871	33.7719298	1.59892717	3.308572e-215	31.310973
P.LEX.CAT=P.LEX.CAT_N	100.00000000	25.6578947	1.20693212	7.701209e-163	27.193744
DAT.FUNC=DM	50.82508251	33.7719298	3.12564473	1.341978e-129	24.220843
P.LEX.CAT=P.LEX.CAT_ADJ	17.55607115	49.7807018	13.33814731	2.021356e-83	19.350433
O.INDIV=O.irr	6.56398445	100.0000000	71.66288426	1.271758e-68	17.506786
O.DEF=O.def.irr	6.56398445	100.0000000	71.66288426	1.271758e-68	17.506786

DAT.FUNC=EXP	11.34246575	45.3947368	18.82607799	1.864673e-40	13.316155
P.TYPE=P.TYPE_PROPERTY	16.60649819	20.1754386	5.71487518	6.770251e-28	10.948305
P.LEX.CAT=P.LEX.CAT_ADV	45.94594595	7.4561404	0.76335878	5.579464e-26	10.541180
P.LEX.CAT=P.LEX.CAT_INTERROGATIVE	100.00000000	2.8508772	0.13410357	4.683001e-18	8.660840
P.LEX.CAT=P.LEX.CAT_D.M	100.00000000	1.9736842	0.09284093	1.045366e-12	7.124398
ELLIPSIS=ellip.yes	15.13157895	10.0877193	3.13596039	1.444781e-12	7.079689
P.TYPE=P.TYPE_SB	14.32835821	10.5263158	3.45574582	3.562840e-12	6.953519
DAT.FUNC=EVAL	12.92682927	11.6228070	4.22942026	1.494375e-11	6.748446
MODE=irrealis	5.43417367	85.0877193	73.65380648	2.139320e-09	5.986859
P.LEX.CAT=P.LEX.CAT_P	100.00000000	1.0964912	0.05157830	2.255287e-07	5.176962
P.TYPE=P.TYPE_SD	9.64749536	11.4035088	5.56014029	6.884314e-07	4.964599
P.VOICE=active	4.82948528	100.0000000	97.40045389	4.525670e-06	4.585654
AFFIRM=negative	8.05755396	12.2807018	7.16938312	6.556552e-05	3.991829
P.TYPE=P.TYPE_VALUE	8.60927152	8.5526316	4.67299360	2.591361e-04	3.653058
DAT.PRSN=scnd	5.56222489	30.4824561	25.77883227	2.055250e-02	2.316106
P.TYPE=P.TYPE_SA	14.28571429	0.8771930	0.28883846	4.946590e-02	1.964554
O.INDIV=O.high	0.00000000	0.0000000	0.71178048	3.555968e-02	-2.101927
DAT.PRSN=thrd	4.00782014	26.9736842	31.65875799	2.610530e-02	-2.224642
V.BIN=V.BIN_HUFAL	0.00000000	0.0000000	0.80462142	2.297038e-02	-2.273927
NO.D=NO.D_A	0.00000000	0.0000000	1.34103569	1.824017e-03	-3.117484
DAT.FUNC=POSSDAT	0.89285714	0.4385965	2.31070765	1.596628e-03	-3.156522
V.BIN=V.BIN_NIFAL	1.85185185	2.1929825	5.57045595	3.619598e-04	-3.566370
AFFIRM=affirmative	4.44493833	87.7192982	92.83061688	6.556552e-05	-3.991829
P.VOICE=pass/mid	0.00000000	0.0000000	2.59954611	4.525670e-06	-4.585654
MODE=realis	2.66249021	14.9122807	26.34619352	2.139320e-09	-5.986859
NO.D=NO.D_V	1.16591928	2.8508772	11.50195998	6.507954e-12	-6.868063
ELLIPSIS=ellip.no	4.36634718	89.9122807	96.86403961	1.444781e-12	-7.079689
P.TYPE=P.TYPE_SC	0.89730808	1.9736842	10.34660615	1.115521e-12	-7.115445
DAT.FUNC=HEP	1.60753880	6.3596491	18.60944914	1.517719e-14	-7.686029
O.DEF=O.def.yes	0.00000000	0.0000000	6.92180730	2.737694e-15	-7.902316

```
S.AGE.INDIV=S.mid                     0.24479804   0.4385965  8.42789354  1.410067e-15  -7.984576
V.BIN=V.BIN_PIEL                      0.11402509   0.2192982  9.04683309  2.756173e-18  -8.721057
O.INDIV=O.mid-high                    0.00000000   0.0000000  8.44852486  1.191012e-18  -8.815548
V.BIN=V.BIN_HIFIL                     1.12970711   5.9210526 24.65442542  3.172375e-27 -10.807488
NO.OF.ARGUMENTS=two                   1.16429495   7.8947368 31.89601816  2.966066e-36 -12.573165
DAT.FUNC=REC                          0.06090134   0.2192982 16.93831236  1.850805e-36 -12.610388
P.TYPE=P.TYPE_PB                      1.35832654  10.9649123 37.97194141  4.834309e-40 -13.244821
O.INDIV=O.mid-low                     0.00000000   0.0000000 18.70229008  7.720854e-43 -13.719865
S.AGE.INDIV=S.high                    0.04701457   0.2192982 21.94140706  5.368839e-49 -14.712376
O.DEF=O.def.no                        0.00000000   0.0000000 21.41530844  9.132830e-50 -14.831745
P.TYPE=P.TYPE_PA                      0.27155465   1.7543860 30.38993192  1.506017e-60 -16.414506
NO.D=NO.D_CL                          0.06851662   0.4385965 30.11140912  2.090031e-69 -17.609303
DAT.FUNC=ADD                          0.15253203   1.0964912 33.81473076  5.221608e-75 -18.325068
V.BIN=V.BIN_KAL                       0.23940627   2.1929825 43.08850836  7.238739e-97 -20.885602
NO.D=NO.D_NP                          0.18832392   1.9736842 49.29853518 2.311084e-121 -23.426208
S.AGE.INDIV=S.low                     0.19179133   2.1929825 53.78584692 3.637677e-138 -25.020724
NO.OF.ARGUMENTS=three                 0.00000000   0.0000000 63.42067258 1.862987e-208 -30.811063
P.LEX.CAT=P.LEX.CAT_TRANS.V           0.02882260   0.4385965 71.58035898 1.624765e-256 -34.211612
PRED.FIRST=predNotFirst               0.07651109   1.3157895 80.89539922  0.000000e+00       -Inf


attr(,"class")
[1] "catdes" "list "
```

ארבעה מבנים אלו מייצגים ארבע קבוצות של מופעי-דאטיב בקורפוס, כאשר כל קבוצה מאופיינת בדפוס שימוש שונה. המחקר מראה כי מתוך נקודת מבט המביאה בחשבון את תנאי השימוש של הדאטיב, ישנה חשיבות מועטה לשיפוט הסובייקטיבי של תפקיד-המשתתף הספציפי. כלומר, החשיבות בהבדלים בין סוגי דאטיבים שונים נעוצה בהבדלים בין תיאורים שונים של המציאות (ברמת מבנה פרופיל השיח), ולאו-דווקא בין תפקידים סמנטיים (ברמת המשתתף הדאטיבי). בעבודה זו אני מראה כי מבני פרופיל השיח מכנסים תחתם קבוצות של תפקידים סמנטיים הקשורים כולם לאותו תיאור מציאות. מבני פרופיל שיח אלו הם הבסיס להשוואה בין היקרויות שונות של הדאטיב, כמו-גם לפרשנות הדאטיב.

או שיחי). כלומר, ההסברים המוצעים בספרות לתפקידי הדאטיב מוגבלים לטבעם של היחסים בין רפרנט הדאטיב והמושא הישיר, למשל, או לשאלת הקשר בין מושפעות ושייכות (בדיונים בנושא דאטיב השייכות), אם לציין דוגמא נוספת. לעומת זאת, המחקר הנוכחי מתאר ומסביר את תפקידי הדאטיב בהקשר רחב יותר, ומושא המחקר אינו מורפמת הדאטיב, כי-אם מבנים דאטיביים. גישה זו לתיאור ופירוש הדאטיב נובעת מתוך תפיסה בלשנית מבוססת-שימוש. בעזרת התבוננות במספר רב של גורמים בו-זמנית, ממקורות שונים של מידע (לשוני וחוץ-לשוני), מחקר זה מפתח סולם מושפעות אשר על-פיו מפורשים מבני הדאטיב בעברית. התייחסות זו למבני הדאטיב בעברית מבוססת על מחקר קורפוס של שפה דבורה תוך שימוש בסטטיסטיקה מרובת-משתנים ובכלי ניתוח נתונים.

בסיס הנתונים לעבודה זו הוא קורפוס של עברית דבורה בן כ-1,760,000 מילים. הקורפוס הוא למעשה אוסף תמלולים של 198 פגישות של וועדות הכנסת השונות, בהן ניתן למצוא רבדים ומשלבים שונים של השפה העברית, גבוהים ונמוכים. בקורפוס נמצאו 16,757 מופעים של דאטיב פרונומינלי (ל-כינוי גוף), כאשר כל מופע כזה נותח וקודד באופן ידני על פני שבעה עשר פרמטרים.

הניתוח המוצע במחקר הנוכחי מבוסס על סטטיסטיקה של חקר נתונים רבי-משתנים ובייחוד על שני הכלים הבאים: Multiple Correspondence Analysis ו- Hierarchical Clustering on Principal Components. Multiple Correspondence Analysis היא טכניקה אשר חושפת קשרים מבוססי שכיחות בבסיס הנתונים, ומאפשרת ייצוג ויזואלי של קשרים אלה על מפה דו-מימדית. באופן זה, למשל, צורות מקושרות יהיו קרובות יותר על המפה מאשר צורות לא מקושרות או מקושרות פחות. Hierarchical Clustering on Principal Components הוא כלי סטטיסטי שנועד לחלוקת נתונים לכדי קבוצות של מופעים דומים. זהו כלי משלים לטכניקה שהוזכרה לעיל, אשר משמש לייצוג גרפי של דמיון או מתאם בין הנתונים ולהגדרה סטטיסטית של קבוצות של מופעים בתוך נתונים.

המחקר הנוכחי מציע קונספט תיאורטי חדש, ומגדיר דפוס שימוש כמבנה של פרופיל שיח: זיווג קונבנציונאלי של מקורות רבים של מידע עם תיאור מסוים של המציאות. מתוך גישה המניחה ייצוגים קוגנטיביים מבוססי-אקזמפלרים, ארבעה מבני פרופיל שיח עולים מן הנתונים:

1. מבנה פרופיל שיח טרנזיטיבי מורחב
2. מבנה פרופיל שיח נקודת הסיום האנושית
3. מבנה פרופיל שיח אינטרנזיטיבי מורחב
4. מבנה פרופיל שיח "נקודת הייחוס"

תקציר

בתפקידו כיחסה דקדוקית, הדאטיב מתאר סוג מסוים של יחסים בין מצב עניינים ורפרנט (בעברית כמו גם בשפות אחרות). ניתן לתאר זאת ככלי אשר השפה מספקת לדובריה על מנת לצייר את הקשר בין סיטואציה לישות בעולם. התפקיד התחבירי אשר משוייך אל הדאטיב, בדרך-כלל, הוא המושא העקיף. הדאטיב מסמן יחסים עקיפים, אשר יכולים לבוא לידי ביטוי בצורת משתתף אשר מושפע חלקית או באופן עקיף מן הסיטואציה, או בצורת משתתף משני, לא אינהרנטי למצב העניינים. סוג כזה של יחסים עקיפים מקושר בדרך-כלל למעורבות קוגנטיבית בסיטואציה (לעומת מעורבות פיזית). תכונה זו של הדאטיב ייחודית לסימון משתתף-מושפע אנושי. באופן גורף, ולאורך שפות שונות, משתתפים-מושפעים אנושיים מסומנים באופן שונה ממשתתפים-מושפעים לא אנושיים (Haspelmath, 2001). כלומר, בניגוד למשתתפים-מושפעים לא אנושיים בהכרח, משתתף-מושפע אנושי מסומן כישות שלא עוברת שינוי מצב מלא.

ישנם סוגים שונים של יחסים עקיפים, וכולם מסומנים על ידי הדאטיב בעברית. תפקיד המקבל, למשל, הוא משתתף עקיף באירוע של הזזת חפץ ממקום למקום (אירוע תנועה טרנזיטיבי). משתתפת שמעריכה סיטואציה מסוימת, גם היא משתתפת עקיפה בסיטואציה עצמה. הסימון המשותף של סוגי יחסים שונים מעלה שתי שאלות. ראשית, מהו טיבו של הקשר בין סוגי היחסים השונים (וראו בונה ובר-אשר סיגל (2014) בעניין זה). שנית, כיצד ניתן להבחין בין סוגי הדאטיב השונים. כלומר, כיצד הנמענת, בשיח טבעי, מפרשת כראוי פסוקית דאטיבית, או מהי האסטרטגיה בה היא נוקטת על מנת להבין את מבעו של הדובר ולפרש נכון את התפקיד אותו נושא המשתתף המסומן על ידי הדאטיב. המחקר הנוכחי מציע תשובות לשתי השאלות, בטענה שישנם ארבעה תפקידי-דאטיב בסיסיים בעברית, כאשר כל אחד מהם קשור לדפוס-שימוש ייחודי, וכולם קשורים אחד לשני על בסיס סולם של מושפעות. דפוס השימוש (ובעיקר דרגת הטרנזיטיביות והסובייקטיביות של הפסוקית) הוא אשר מכוון את הנמענת לפרשנות הנכונה.

הנחה זו לא בהכרח מתאימה לניתוחים מוקדמים יותר של הדאטיב, בעברית או בשפות אחרות. באופן מסורתי, תפקידי הדאטיב מתוארים ומוסברים ללא תלות בהקשר בו הם מופעים (הקשר תחבירי

עבודה זו נעשתה בהדרכת

פרופסור מירה אריאל

אוניברסיטת תל-אביב
הפקולטה למדעי הרוח ע"ש לסטר וסאלי אנטין
בית הספר למדעי התרבות ע"ש שירלי ולסלי פורטר

מיפוי מבני הדאטיב בעברית

חיבור לשם קבלת התואר "דוקטור לפילוסופיה"

מאת

אליצור דטנר

הוגש לסנאט של אוניברסיטת תל-אביב

אוגוסט, 2014

אוניברסיטת תל-אביב
הפקולטה למדעי הרוח ע"ש לסטר וסאלי אנטין
בית הספר למדעי התרבות ע"ש שירלי ולסלי פורטר

מיפוי מבני הדאטיב בעברית

חיבור לשם קבלת התואר "דוקטור לפילוסופיה"

מאת

אליצור דטנר

הוגש לסנאט של אוניברסיטת תל-אביב

אוגוסט, 2014