

# PhD Thesis Proposal v10

Utku Can KUNTER

June 12, 2019

## Contents

1	Introduction	2
2	Problem Definition	4
2.1	Syntax vs. Morphology . . . . .	4
2.2	Inflection vs. Derivation . . . . .	4
2.3	Which Subset of Derivation? . . . . .	5
2.4	Generative vs. Constraint-Based Grammars . . . . .	6
3	State of the Art	7
3.1	Morphology at Large . . . . .	7
3.2	Some Particular Aspects of Morphology . . . . .	9
3.3	Representing Grammatical Categories . . . . .	10
3.4	Turkish Morphology . . . . .	10
3.5	CCG . . . . .	11
4	Shape of the Problem	13
4.1	Representations of Major Grammatical Categories . . . . .	13
4.2	An Inventory of Turkish Affixes . . . . .	16
4.3	A Glance at Inflectional Processes . . . . .	17
5	Preliminary Work	21
5.1	Deriving from Nouns . . . . .	21
5.2	Deriving from Verbs . . . . .	22
5.3	Deriving from Adjectives . . . . .	26
5.4	Deriving from Adverbs . . . . .	28
5.5	Revisiting Inflection vs. Derivation . . . . .	29
5.6	A Deeper Glance at Inflection: Marking Indefiniteness in Turkish . . . . .	30
	References	39

# 1 Introduction

In preparing this proposal, our aim has been to discover a topic within linguistics, that can be studied with a focus on cognitive science and where we can make a difference, however small.

While there has been a massive amount of research in the fields of syntax, and inflectional morphology, it appears derivational morphology has been quite understudied. Several possible reasons for this immediately come to mind. First of all, syntax, more precisely syntax at the individual word level, has been under the spotlight since the early years of generative grammar. Working at word level is the most natural approach to the problems of generative grammar, as the segmentation is already there and the number of elements to be accounted for is at a minimum for a given sentence. Of course, further analyses demonstrated how a word-level investigation fails to account for many phenomena and that a system of regularities also exist below the word level. The next logical step was to come up with a theory of inflection, as the morphological regularities relevant to syntax are (arguably always) the product of inflectional processes. Therefore, inflection has also been studied a lot. As a natural consequence of the way the research field evolved, derivational morphology had to wait the longest to step into the spotlight.

A second possible reason is the inherent complexity of the problems related to derivational morphology. Derivational processes contribute a wide variety of semantic content, must satisfy numerous semantic constraints before they can be applied and can wildly change the argument structure of the host. Many other challenging aspects of derivation could be added to the list. On the other hand, inflectional processes occur within the realm of syntax, which is better understood, their application is often (perhaps always) mandatory and their semantic contribution is relatively standardized. All in all, inflection seems to be the easier part of morphology to deal with.

The general research tendencies in the field are reflected in the collection of studies on Turkish. Derivational morphology has been investigated within the context of descriptive grammar and morphological analysis / disambiguation, but the work on a generative theory of derivation in Turkish, to the best of our knowledge, has not been comprehensive.

In this thesis, our main aim is to help fill this gap in the literature and come up with a theory of derivational morphology in Turkish. Our focus is limited to Turkish, because the variety of semantic possibilities would otherwise be endless. Even a comprehensive work on Turkish derivational morphology could prove to be too much to handle, but this is not an excuse for ignoring this research area altogether.

Our hope is to shed a little light on the intricacies of Turkish derivational morphology. By emphasizing the differences of derivational processes from syntactic processes, we will first narrow down the research question. Then, we will try to explain the regularities in these processes with a perspective of generative grammar.

Taking up this challenge, we need a tool with an adequate expressive power; flexible enough to allow a comprehensive analysis of a wide range of phenomena, but restrictive enough to prevent theories from becoming too loose to be considered a natural language grammar. Combinatory Categorical Grammar has over decades accumulated an impressive track record proving its capacity for such an analysis. An important feature of CCG is its handling linguistic processes with an equal regard for syntax and semantics. As derivational processes are more heavily involved with semantics, this proves to be a major advantage of CCG over alternative frameworks that assume a syntax-first approach.

In the following sections, we first determine the boundaries and method of our study. Second, we go over several important works relevant to our problem and draw a general picture of the state of the art. Third, we give a brief overview of the problem space. Finally, we present the results of our preliminary analyses and develop ideas on the possible ways forward.

## 2 Problem Definition

In this section, we try to narrow down the problem by describing our position regarding the relevant dichotomies within linguistics. These include syntax vs. morphology, inflection vs. derivation (vs. construction), derivational affixes with Turkish origin vs. others, and CCG vs. MP.

### 2.1 Syntax vs. Morphology

Possibly one of the most central and most controversial dichotomies in the field is the one concerning syntax and morphology. There have been important works siding with the idea that morphology must be considered a distinct module, such as Aronoff (1994) and Aronoff and Fudeman (2005), while others make the very opposite claim like Lieber (1992). We do not make any claims regarding this dichotomy. On one hand, it really seems morphology distributes itself to different linguistic modules and operations. On the other hand, we feel this does not eliminate the necessity for us to come up with a theory of morphology.

Regarding the mechanism of morphological operations, there are mainly three approaches: item and arrangement (IA), item and process (IP) and word and paradigm (WP). While these approaches are ultimately equivalent in power, they adopt different mechanisms and make different assumptions. IA assumes segmental morphology; it must find ways to represent apparently non-segmental morphology in a segmental manner. In contrast to IA, IP assumes a procedural mechanism, and represents even obviously segmental morphology as processes. WP makes neither assumptions, which is why it seems to be the safest choice. We can say our point of view is closest to WP.

Regarding the structure of morphological operations, an important issue has been raised by Sezer (1991). Sezer explains that syntax is relevant to the thematic structure of a sentence, while morphology is relevant to the argument structure. There are quite important implications of this claim. First, it allows us to draw a more grounded boundary between syntax and morphology. Parallel to this, it also allows us to draw a boundary between inflection and derivation. Thematic structure and argument structure offer us different points of view relevant to our modeling efforts. Of course, these implications depend on the existence of two distinct structures concerning arguments and thematic roles in a phrase, as claimed in Grimshaw (1990) and several others. Our preliminary analyses demonstrate that there is indeed a fundamental difference between how syntactic and morphological processes interact with the syntactic structure. These analyses and comments will be presented in Section 5.

### 2.2 Inflection vs. Derivation

Inflection and derivation are like siblings that are radically different in some ways, but still quite similar in others. This analogy implies it is possible to partition morphology into two non-intersecting sets, but quite often the choice between these sets is a matter of definition. It seems

individual affixes do not occur on a binary scale; they are distributed on a continuous line. There are several alternative places where a demarcation line could be drawn between inflection and derivation, but all such lines would be arbitrary in some sense.

Still the distinction between inflection and derivation helps us by providing shortcuts to refer to groups of affixes that exhibit certain properties. In a general and somewhat imprecise sense, inflection is closer to syntax, is more productive and contributes a well-defined abstract meaning, while derivation is subject to semantic constraints and may change the original concept altogether. These are the traditionally accepted differences between the two sides of morphology. A clear-cut distinction between inflection and derivation may or may not be possible eventually, but initially, we simply claim that morphological processes that are required by syntax are inflection, and the rest are considered derivation. We will take our next steps carefully following Sezer (1991) who claims that these two kinds of processes differ in the ways they interact with the sentential structure.

### 2.3 Which Subset of Derivation?

Derivational morphology as a whole is still an enormous field that precludes a truly comprehensive analysis. Therefore, we restrict our focus to Turkish. As the following sections will show, Turkish involves a rich and diversified derivational morphology. The inventory of affixes given in Bozşahin (2018) is the clearest possible demonstration of derivational capabilities in Turkish language.

A quick glance at this list, however, reveals that some affixes familiar to Turkish speakers are missing. Many affixes are omitted from this list, despite their productive use, due to their non-Turkish origins. Affixes like "na-" from Arabic (i.e. "na-tamam") and "-syon" from French (i.e. "fermant-asyon") are quite productive and they are consciously deployed (and even sometimes removed) by Turkish speakers. Many other affixes, like "mü-" from Arabic (i.e. "mü-dahale") and "-aj" from French (i.e. "arbitr-aj"), cannot be considered productive at all. Words containing these affixes are not analyzed by (most) Turkish speakers, they are kept in the lexicon in their final form. Therefore, there are no derivational processes in Turkish that involve these suffixes, at least for our purposes.

These affixes and their kind present different dynamics than affixes that originate in Turkish. Failing to recognize this distinction, the distinction between original and borrowed derivational processes, would ultimately require a study to cover an enormous list of derivational processes. We avoid this by limiting our investigation to affixes originating in Turkish. Arguably, derivational affixes with Turkish origin are used in a more productive way by Turkish speakers, which make them more valuable in our study.

## 2.4 Generative vs. Constraint-Based Grammars

The standard Chomskian theories of language have been dominant in the literature for decades now. Their prominence came from revolutionary insights into notoriously complex problems about language and the mind itself. However, worthy alternatives appeared in time to challenge their standard status. Constraint-based grammars, which follow a completely different strategy than the Chomskian generative grammars, have been developing rapidly. Among these, CCG has received considerable attention in recent years, due to its adequate expressive power and efficient parsing.

Several languages with interesting properties demonstrate the inadequacy of the syntax-oriented approach typical of Chomskian theories. For instance, split ergative languages like Basque, Georgian and Mayan, violate many of their central hypotheses and force them to adopt ever more intricate designs to account for these seemingly "outlier" languages. Also, they fail to give a satisfactory account of unbounded constructions. (Bounded constructions are more achievable, since they are open to context-free representation.) On the other hand, as Steedman and Bozşahin (n.d.) (hereafter PFTL) demonstrates, split ergative languages do not keep CCG theorists awake at night, and unbounded constructions can be represented quite elegantly in CCG.

With their emphasis squarely on the autonomy of syntax, Chomskian theories set aside semantics and run into great difficulties with phenomena more involved with semantics. Proponents of CCG prefer an alternative strategy by envisioning syntax and semantics simultaneously being "projected from the lexicon" (PFTL). By not ignoring semantics in favor of an autonomous syntax, CCG provides a much more suitable platform to study derivational morphology, which is much more involved with semantics than inflectional morphology and syntax.

### 3 State of the Art

In this section, we briefly describe the current state of the art. First, we list several important studies on morphology in general. Next, we look into some particular aspects of morphology. Third, we try to come up with an adequate set of CCG representations for major grammatical categories. The final two subsections are devoted to important research on Turkish morphology and CCG, respectively.

#### 3.1 Morphology at Large

Grimshaw (1990) puts forward a convincing account of argument structure (a-structure), demonstrating how and why it exists as a distinct layer. Grimshaw (1990) argues that the lexical conceptual structure (lc-structure), which is present in every lexical entry and which is projected directly from the semantic content of the entry, generates a thematic structure ( $\theta$ -structure) and an event structure (e-structure). The  $\theta$ -structure determines the list of arguments that can be licensed by the lexical item and includes an implicit hierarchy between these arguments. On the other hand, the e-structure constitutes an aspectual dimension on top of the  $\theta$ -structure. This aspectual dimension takes priority during the labeling of arguments as subject or object. From the combination of the  $\theta$ -structure and the e-structure yet another structure is born, named the a-structure. The a-structure does not contain any  $\theta$ -marks, but it includes a complete and final hierarchy between all the arguments projected from the lc-structure.

Since the hierarchies from the  $\theta$ -structure and from the e-structure coincide most of the time, the existence of a separate a-structure has not always been obvious. What Grimshaw (1990) does is to put forward a theory that is able to explain the outcomes produced in the presence of conflicting structures. Indeed, analyses of several linguistic phenomena involving psychological verbs and complex event nominalization (as well as many others) show how the two structures do not have to coincide. With clear examples, Grimshaw (1990) demonstrates how her theory explains these phenomena quite elegantly. Hale and Keyser (2002) provides a simpler, beginner-friendly discussion of a theory of argument structure.

Assuming every lexical entry is associated with a thematic structure, and citing several sources for the possibility of an a-structure distinct from the  $\theta$ -structure, Sezer (1991) claims that syntax operates on the thematic structure, while morphology operates on the argument structure. He points out how most derivations involve the addition or suppression of an argument, giving as a typical example passive by-arguments that act both like adjuncts and arguments. This is a powerful claim that touches the heart of the issues we are attempting to investigate. The deepest differences between syntactic and morphological processes may be rooted in their behavior being responsive to different linguistic structures. We expect this claim and arguments accompanying it will be a recurrent theme in our study.

Sezer (1991) also emphasizes the distinction between semantic selection (s-selection) and category selection (c-selection). This distinction may also take an important role in our investigation of the behavior of derivational morphology. One could say s-selection is a more general

kind of restriction and c-selection is a subset of it. While scanning the semantic content of a candidate, s-selection also implicitly takes a position regarding their grammatical categories. On the other hand, c-selection only considers the grammatical categories of candidates. Perhaps inflectional processes are constrained by the principles of c-selection, while derivational processes are subject to a wider set of constraints resulting from s-selection.

Lieber (1992) claims that any processes attributed to morphology can actually be explained within the theory of syntax; succinctly, in her theory, morphology is what we call syntax below the level of  $X^0$ . She argues that compounds are syntactically formed and inflectional processes are basically a part of syntax, but concedes that no one has yet succeeded in deriving the properties of words and sentences from the same basic principles of grammar. We do not care much about this distinction. We are aware that morphology distributes itself to several layers of linguistic modules: Inflection is obviously closer to syntax; some languages keep most relevant content in the lexicon, while some keep it in the grammar. In the end, what we care about the most is being able to represent the entire set of possibilities from syntax and morphology in the simplest possible framework.

Aronoff (1994) argues for the autonomy of a morphological module. He starts with the Separation Hypothesis and gives a history of theories of morphology.

Separation Hypothesis: Morphological operations should be separated from accompanying phonological operations.

Aronoff (1994) demonstrates the existence of a morphomic level, which he claims is a purely morphological level. He distinguishes between inflectional classes and gender, as two autonomous levels and argue that inflectional classes are also purely morphological. He admits morphology is not necessary for a language, but many languages do have elaborate morphology which cannot be adequately explained without a dedicated theory of their morphology. We agree with him on the necessity for a theory of morphology.

Aronoff and Fudeman (2005) provide interesting cases and detailed explanations on both inflectional and derivational morphology from Kujamaat Joola, covering a wide range of possibilities with examples from many languages from around the world. They claim that understanding the morphology of individual languages leads to the understanding of morphology of Language. This is encouraging for us, since we choose to reduce the range of our focus to Turkish derivational morphology. As in Aronoff (1994), they argue for the existence of a distinct linguistic component called morphology, claiming some aspects of morphology cannot be attributed to anything else.

Aronoff and Fudeman (2005) define inflection as the formation of grammatical forms of a single lexeme, uses of which are determined by syntax, but derivation as the creation of one lexeme from another, including compounding. They also discuss the IA-IP distinction and points out the two approaches being equivalent mathematically, provided negative morphemes are allowed in IA. We believe WP provides a more flexible approach than both.



Carstairs-McCarthy (2010) looks for the roots of morphology in the central and pervasive tendency of synonymy avoidance. After going over the conventional arguments on the existence and non-existence of a distinct morphological module, he investigates the possible mechanisms behind the creation of affixes and stem altering operations. Throughout the book, he mostly deals with inflection, but reserves Chapter 7 to derivational processes, where he puts a large emphasis on cliché patterns, collocations, and more generally on lexicalization. In our opinion, the alleged linguistic tendency called synonymy avoidance, although it obviously exists, does not seem strong enough to fuel all the processes listed in the book. Therefore, the abductive argument presented in the book does not seem very convincing.

### 3.2 Some Particular Aspects of Morphology

Universality of Theta Assignment Hypothesis (UTAH) presented in Baker (1988) is quite a central idea in the study of morphological processes. As a consequence of UTAH, the distance of an affix to the stem is determined by its place in the thematic structure and derivational affixes are closer to the stem than inflectional suffixes. This observation proves to be quite an important clue for us. Whenever we examine how an affix interacts with the sentential structure, UTAH provides additional evidence.

Moens and Steedman (1988) provides an outstanding analysis of universal temporal categories and temporal relations between events. The claims made in the paper are not based on purely temporal primitives, but on causation and consequence. Following the classifications and strategies in Moens and Steedman (1988), quite complex event structures may be analyzed in a robust and elegant way. Also, thanks to the framework given in the paper, the vast variety of possibilities from Turkish T/A/M markers can be easily represented in logical forms.

Perhaps Turkish emphatic reduplication (TER) deserves a special mention. The kinds of reduplication used in Turkish (i.e. "tabak mabak", "zaman zaman", "sapsarı") always indicate a change in meaning and are not required by syntax. Thus, they fall into derivational morphology. One can also say, however, reduplication should be considered a morphophonological process, since there is no allomorphy. So, there are reasons to include Turkish reduplication in our study, and there are reasons to exclude it. For now, we leave the choice to a later stage and stop at a shallow introduction to arguably the most interesting kind of reduplication in Turkish, the emphatic reduplication.

Wedel (1999) and Kılıç and Bozşahin (2013) present important findings on TER; the former looking into the phonological aspects of the phenomenon, while the latter prefers a data-driven approach. Wedel (1999) claims that reduplicative forms used for putting emphasis on adjectives are somewhat productive. Working within the framework of Optimality Theory, he shows that the schema for creating new forms is available to Turkish speakers.

Kılıç and Bozşahin (2013), on the other hand, show that in addition to the phonological constraints, speakers show two preferences in their linker type selection. First, they try to dissimilate the linker type from frequent consonant co-occurrences and word endings. Second, they prefer the linker type to be dissimilar from the existing root in their choice of linker type. Therefore, they claim, TER is a morpholexical process that depends on "a global lexical knowledge for selecting an appropriate linker".

### 3.3 Representing Grammatical Categories

Achieving an adequate representation of each major grammatical category is crucial, as the success of morphological operations depends on the host's capabilities for receiving new categories and semantics from affixes. Even at the preliminary stage of our analysis, we needed to make use of complicated logical forms for host grammatical categories and subcategories.

While noun phrases seem to be a simple, even trivial class in terms of their logical form representation, Grimshaw (1990) Chapter 3 presents quite detailed arguments against this default position. Nevertheless, at this initial stage, we will use the simplest possible representation for noun phrases, and gradually move onto more complex representations.

At present, our representation of verb forms is similar to PFTL's. In time, following important works such as Van-Valin-Jr. (2006), we will try to achieve an even deeper analysis. PFTL makes important claims shaping our strategy for building a CCG. Their assumptions in Chapter 4 regarding verbs and verbal morphology guide us in Sections 4.1 and 5.2. In these sections, we attempt to find a way to represent verbs that will accommodate all the morphological processes we will have to deal with.

Our review of the adjectival semantics literature has been relatively more complete. We cite Paoli (1999), Paradis (2001), Kennedy and McNally (2005) and Kennedy (2007) in Section 4.1 to explain the reasons behind our decisions in the construction of logical forms for adjectives. For adverbs, the literature has not been as generous, but for now, we modify and reuse the principles from adjectival semantics.

### 3.4 Turkish Morphology

Sezer (1991) is an outstanding work that makes important claims regarding Turkish syntax. Its most relevant and central claim for the present work is that syntax interacts with the  $\theta$ -structure, while morphology interacts with the a-structure. The reasons for this claim and its consequences are discussed at length in Section 2 and Subsection 3.1.

Oflazer et al. (1995) gives "an outline of Turkish morphology", complete with a list of morphophonemic processes, an affix inventory and finite-state machines for nominal and verbal morphotactics. Bozşahin (2018) builds on the affix inventory presented in this work. He creates a list of primitive binary concepts and analyzes each affix in the inventory using these concepts.

Subections 2.3, 4.2 and 5.2 are largely based on this paper.

Göksel and Kerslake (2005) provides a truly comprehensive grammar of Turkish. This extremely detailed and perfectly organized book is an invaluable resource for beginners in Turkish grammar. Erguvanlı-Taylan (2001) provides an even deeper analysis of Turkish verbs. Looking at various aspects of verbs, Erguvanlı-Taylan (2001) gives important insights to the reader.

### 3.5 CCG

Bozşahin (2002) argues that setting aside inflectional morphology as a word-internal process and "designating words as minimal units of the lexicon" is too restrictive. Instead, in this study, inflectional morphology and syntax are integrated within the framework of CCG and a morphemic lexicon. The ideas put forward in this paper are in harmony with our approach. In fact, representation of grammatical categories and affixes in this paper form the basis of many representation strategies in the present work.

Steedman and Baldridge (2011) is a complete introductory level discussion of CCG's capabilities and basic grammar-building strategies. It is a primary resource in the matters related to CCG.

Bozşahin (2017) is the manual for CCGab, the computational tool on which all examples in this study are tested. CCGlab is a blessing, to say the least, making rapid prototyping possible for complex CCG rule sets. With a reliable and easy-to-use engine, it implements the CKY parser and reduces the time required for testing by several orders of magnitude.

PFTL launches a full-scale investigation of CCG's capabilities, with demonstrative examples from numerous languages. While most of the book is relevant to our purposes, especially the Chapters 4 and 16 deserve a special mention here.

Chapter 4 of PFTL looks into the categorial lexicon, establishing a framework and a general strategy for determining the syntactic categories and the logical forms of lexical items. The claim by Sezer (1991) that syntactic processes and derivational processes handle argument structure differently is reflected in CCG in quite an elegant way. (Steedman and Baldridge (2011) gives a complete list of modes of application on p.9.)

#### (1) Application rules

- (a) Double slash: -ci (as in "muhasebe-ci"):  $N \backslash N: \lambda x \lambda a. deals\_with' x a$
- (b) Single slash: -di (as in "gel-di"):  $(S \backslash NP_{NOM}) \backslash V: \lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$

They use single slashes for syntactic operations and double slashes for derivational morphology. Double slashes ensure that only the application operation is allowed and both the input to and the output from the application must be lexical. These are powerful restrictions that hint at a

fundamental difference between the nature of the operations that are restricted by them and the ones that are not. Indeed, double slashes are always present in derivational processes but never in inflectional processes. Could their role be used to explain the difference between thematic structure and argument structure? Either way, CCG provides us an adequate platform for such an analysis.

Chapter 16 of the same work mainly deals with relativization and coordination in Turkish. These two are significant in that they both allow for unbounded constructions, which are quite tricky to explain in any theory, but can be elegantly derived in CCG.

The discussion on relativization benefits from an important observation made by George and Kornfilt (1981). George and Kornfilt (1981) suggest that the genitive markers on the relative clause subjects are not actually genitive markers but agreement markers. They contribute a sense of finiteness to the clause, agreeing with the verb. Agreement markers on the verb are also homophones of the possessive marker, but there is no sense of possession in the construction. This convincing claim is quite significant, because it greatly simplifies the CCG representation of relative clauses.

## 4 Shape of the Problem

A study of derivational morphology must take place within a network of other linguistic modules that are themselves still under investigation. Syntax and semantics are always relevant in the decisions one has to make regarding the best representation of affixes' syntactic categories and semantic content. Inflectional morphology is also relevant due to the potential clues hidden in the intersection of inflection and derivation, and in the different behavior of the two halves of morphology with respect to argument structure. Needless to say, phonology appears on stage from time to time, but is mostly ignored due to its effects being deemed too close to the surface.

In this context, it would be inadequate to wander forever in the realm of derivational morphology. While we believe derivational morphology is a system on its own, its interactions with other systems must be part of the investigation. Studying it in isolation would fail to recognize its functions within the wider system and miss the clues hidden on the thick and blurred boundaries between linguistic modules. In this section, we try to take a low resolution picture of the wider system, and set the stage for our investigation.

First, we explain our initial choices in representing major grammatical categories. Second, we take a general look on Turkish morphological processes, based on an inventory of Turkish affixes. Then, we review inflectional processes and try to identify properties that distinguish them from derivational affixes.

### 4.1 Representations of Major Grammatical Categories

We have already pointed out that derivational affixes are capable of changing the grammatical category of an input. This adds a new layer of complexity to our task. Before starting an investigation of derivation, we must decide how to represent the main classes of lexical items. These representations should be carefully constructed so as to accommodate a wide range of derivational operations.

We have four major grammatical categories to account for: nouns, verbs, adjectives and adverbs. We leave propositions aside for the moment.

Nouns are represented in the simplest form. Their syntactic category is N. In their bare form, nouns are not taken to be cased and must get a case to assume their role within a sentence, as prescribed by the case filter (Cowper (1992)). Named entities are no exception; they may take any case marker, and they must take case. Features would be necessary to track information required for person and number agreement, but this is not an immediate concern.

#### (2) CCG representation of nouns

- (a)  $ev : N : \text{home}'$
- (b)  $Ankara : N : \text{Ankara}'$

Verbal categories are not too complicated either. PFTL Chapter 4 presents important assumptions regarding the representation of predicates in CCG. The following is the list of these assumptions:

(3) PFTL assumptions on verbal representation

- (a) Bare infinitives (or root forms) are functions onto V.
- (b) Non-bare infinitives are functions onto VP.
- (c) Finiteness is a function onto S|NP for some |.
- (d) Aspect and event modalities are functions from VP and S|NP for some |.

They add that verb roots are assumed "to be the determiners of thematic structure and argument structure for predicates".

Verbs in their bare form produce a V after their internal arguments are fulfilled. They must be promoted by T/A/M and agreement markers to the sentence level (S|NP for some |) before they can be joined with a subject and produce a complete sentence. This requirement is in line with the syntactic structures explained in Cowper (1992).

Verbs may be intransitive, transitive or ditransitive. A verb's arity is reflected in both its syntactic category and its logical form, with an appropriate number of bound variables, denoted by  $x$ ,  $y$ ,  $z$  and so on. The variable  $a$  in the logical form stands for the subject NP. Following Moens and Steedman (1988) we include bound variables  $t$  and  $w$  for representing event time and world. These variables are indispensable in representing the semantic contribution of T/A/M markers as well as some derivational markers.

(4) CCG representation of verbs

- (a) gel:  $V: \lambda a \lambda t \lambda w. come' at w$
- (b) gör:  $V \backslash NP_{ACC}: \lambda x \lambda a \lambda t \lambda w. see' x at w$

Our discussion of verbs have been informed mainly by PFTL so far, but in the next steps, as a deeper understanding of verbal semantics becomes necessary, Van-Valin-Jr. (2006) will be referred to a lot more frequently. In this book chapter, a comprehensive analysis of verbal semantics is presented, including event semantics and examples from multiple languages.

Each adjective expects a noun and acts like a predicate. Its logical form must contain a bound variable that will be fulfilled by the noun. If we do this in the simplest way and do not use any other bound variables, receiving the expected noun completes the term. As a result, we do not have any bound variables left to link the entity denoted by the noun to the remainder of the sentence. We prefer keeping a second bound variable to represent the entity.

Adjectives come in two shapes. Many adjectives are gradable in the sense that the noun modified by one is only deemed worthy if it has the property to some predefined extent. For

example, for a car to be fast, it must be faster than some standard level of fastness. This level is stored in each speaker's context and somehow implicitly conveyed to the hearer. There are also adjectives that are not gradable, or at least their scale is not continuous, but binary. Paoli (1999) calls the former kind relative adjectives, and the latter absolute adjectives. In the CCG representation attempts below,  $x$  is a placeholder for the noun to be modified;  $a$  stands for the generic entity that is both modified by the adjective and which is an instance of the noun;  $d$  is the degree at which the sense of adjective's is observed and  $l$  is the level of acceptance beyond which the adjective's use is justified. For non-gradable adjectives, we do not need  $d$  or  $l$ .

(5) CCG representation of adjectives

- (a) hızlı:  $N/N: \lambda x \lambda a \lambda d \lambda l. (quick' (a' d) a) \wedge (d > l) \wedge (xa)$
- (b) anayasal:  $N/N: \lambda x \lambda a. (constitutional' a) \wedge (xa)$

Paoli (1999) makes an important warning at this point, regarding the representation of relative adjectives. He points out that the standard level that permits or prevents the use of a given adjective depends on the noun modified by it. For instance, the standard level of fastness for a human should be different than the standard level used to judge a car. Thus, one can say, this dependency must be represented in the logical form of the adjective. In the article, this point is demonstrated with the following example:

(6) Semantics of gradable adjectives

- (a) Dumbo is a small elephant.
- (b) Dumbo is small and Dumbo is an elephant.

We need to modify the logical form in (5) (a) to make the level of acceptance  $l$  dependent on the noun. But how do we do this? Are there really just two kinds of adjectives, or should we look deeper into the semantics of adjectives and try to accommodate several more kinds? Paradis (2001) gives a nice schema for the kinds of adjectives; including the gradable-non-gradable dichotomy, as well as considering oppositeness and boundedness. Kennedy and McNally (2005) and Kennedy (2007) create a massive inventory of possibilities for representing adjectival semantics, supported with arguments and counterarguments for each strategy.

Recognizing that adjectival semantics is an ocean of challenges on its own, we choose to remain at a shallower level of analysis for now, keeping our focus on morphology. As we progress in our study of affixes, a deeper analysis of adjectives will undoubtedly become necessary, since each class of adjectives may react differently to each class of affixes. Due to the wide variety of semantic constraints on derivational morphology, an adjective may allow the application of one affix, while disallowing another, or its argument structure may be altered in a different way than previously expected. We will need to reflect these constraints on our representations by recognizing the variety of adjective classes and tuning our representations to accommodate them. Our initial analysis will involve the gradable-nongradable dichotomy, and other levels of analysis will be added later, as they become necessary. Trying to cover them all as soon as possible

would shift our focus from the main issues of interest, and could increase the complexity of the task beyond manageable levels.

Adverbs specify when, how or why an action takes place. Time adverbs require us to represent the time relation explicitly in the logical form. A specification of manner can be conveyed in the way adjectives modify nouns. Just like how the existence of some item  $a$  is postulated by an adjective, which is then modified both by the adjective and the noun itself, an adverb could postulate the existence of an action  $a$ , first modified by the adverb and then the action along with all its arguments. For time adverbials, we sometimes need extra bound variables to represent the some information from context. For instance, in (7) (a),  $t_d$  denotes the amount of time used to judge if a time difference is small enough for the adverb "demin" to be used. Reasons could be added inside the matrix verb's term as an adjunct, until a better way is found. Since the adverbs of this kind require propositions, we do not delve into their analysis immediately.

(7) CCG representation of adverbs

- (a) demin:  $V/V: \lambda p \lambda a \lambda t \lambda w \lambda t_d. (patw) \wedge (t < t_0) \wedge (t_0 - t < t_d)$
- (b) hızlıca:  $V/V: \lambda p \lambda x \lambda a \lambda d \lambda l. (quick' (at' d) a) \wedge (d > l) \wedge ((px) a)$

Of course, this is also a preliminary analysis. Gradability and other dichotomies should again be present in a full investigation of adverbs. As well as modifying verbs, adverbs may modify adjectives and other adverbs, adding another level of complexity to their representation. Degree adverbials should perhaps be considered a wholly different class, while time adverbials can be assumed to only modify verbs etc.

Such constraints seem to be grounded in the way lexical items' argument structures fit with each other. For instance, adverbs of time, expect a grammatical category that carries a time variable, and that category happens to be the verb. So, it is not that time adverbials expect a verb, verbs happen to be the only grammatical category that fulfills the criteria for being modified by an adverb. This line of thinking will take a more central place in the discussions on morphology. Not to repeat the reasons given above, suffice it to say that these levels of detail will be added to our study as their inclusion becomes an asset than a liability.

## 4.2 An Inventory of Turkish Affixes

What better place is there to start an analysis of morphology than a complete inventory of affixes? Bozşahin (2018) makes a bold attempt at determining the binary concepts underlying Turkish morphological processes. Based on Oflazer et al. (1995), his inventory of affixes is claimed to be comprehensive, but admittedly preliminary. The idea is to come up with a conceptual analysis of these processes to provide a reliable basis for further research on Turkish morphology, syntax and semantics. A few of these binary concepts (out of a total 24) are CAUS (causative), FACIL (facilitative), GRAD (gradable), PRED (predicative) and STAT (stative).



Bozşahin (2018) uses a coding scheme to distinguish each class of affixes, including the final grammatical category, the source grammatical category, and the type of process (inflectional or derivational). He adds a representative allomorph (for derivations) or a function description (for inflections) to this term and assigns each affix a unique code:

(8) Coding scheme in Bozşahin (2018)

- (a) NNI\_PLU: Inflection on a noun, producing a noun, making it plural
- (b) VVI\_REFX: Inflection on a verb, producing a verb, making it reflexive
- (c) VJD\_AL: Derivation on an adjective, producing a verb, using a suffix of the form AL
- (d) AAD\_CEK: Derivation on an adverb, producing an adverb, using a suffix of the form CEK

Inflectional affixes in Bozşahin (2018) never change the grammatical category of their input item, while derivational affixes may or may not make such changes. Arguably, the most interesting derivations occur between predicates and non-predicates, since such derivations require the affix to make substantial changes on the argument structure. While inflectional affixes can be studied in just a few groups, the semantic contribution of derivational affixes prevents us from working at a similar level of complexity.

Bozşahin (2018) gives us a general view of the affixal inventory. The following tables show the distribution of affixes with respect to their function. Bozşahin (2018) makes the adjective-noun distinction, however adds that the distinction is not as clear as the one between verbs and nouns etc. and does not list any zero affixes. We believe recognizing a zero affix for converting adjectives to nouns is inevitable. The logical forms of adjectives are quite different than nouns, and a conversion must be carried out by a zero affix (or perhaps a unary rule, if we are following IP) before any nominal affix can be applied. This is why the adjective to noun count in Table 2 is marked with an asterisk.

Source / Result	Noun	Verb
Noun	19	0
Verb	0	41

Table 1: Distribution of Inflectional Affixes

The following subsection provides possible representations of some selected inflectional affixes in CCG. Derivational affixes will be attempted in Section 5. Based on these, we will try to discover regularities.

### 4.3 A Glance at Inflectional Processes

Nominal inflection in Turkish is fairly simple, at least compared to verbal inflection. There are only three main classes of markers: plural, possessive and case. They all carry out their tasks at

Source / Result	Noun	Verb	Adjective	Adverb
Noun	9	6	10	6
Verb	27	6	20	2
Adjective	2*	6	6	3
Adverb	0	0	0	1

Table 2: Distribution of Derivational Affixes

the syntactic level (We will return to ”-ki” later in this subsection.)

(9) Plural marker

(a) -ler:  $N \setminus N: \lambda x. multiple' x$

(10) Case markers

(a) -Ø:  $NP_{\text{NOM}} \setminus N: \lambda x. x$

(b) -i:  $NP_{\text{ACC}} \setminus N: \lambda x. x$

Please notice the use of  $\setminus$  in the syntactic categories. These affixes must be more flexible in their application, as we may have to postpone their application until after some modifiers are applied or some constructions have been set up.

(9) (a) representation of a plural marker is admittedly quite simple and open to improvement. Perhaps, like many inflectional operations, the change that takes place in the semantics should just be reflected on the syntactic category, with the help of a feature. In fact, we must perhaps use a number feature in preparation of verbal agreement inflection. We did not go into that level of detail at this stage. (10) case markers demonstrate that strategy by employing a feature indicating case, but no change in the logical form.

If we accept there is a zero marker indicating nominative case, which one might say leads to a more orderly and simple grammar, case marking is obligatory. Case marking ensures that each noun phrase assumes its intended place within the thematic structure. Since it carries out such an important task, its being obligatory does not sound surprising. The plural marker carries out a comparable task, ensuring agreement in number. It would not be surprising to imagine an obligatory number marker that distinguishes between singular and plural nouns. As in most languages, the ”default” forms of nouns, namely singular and nominative, are not marked in Turkish. It might give us a clearer field if we assume inflectional operations are always obligatory. Just like how a zero marker makes a noun nominative, many other zero markers may be present to mark default forms. If we refuse to give the default forms a different status than other forms, we might be able to see more regularities in the data.

We believe the relative marker ”-ki”, which is mentioned under the NNI group in Bozşahin (2018) should be considered an exception.

(11) Relative marker -ki

- (a) ev -de -ki ses
- (b) araba -da -ki -ler -in -ki
- (c) sen -in -ki

”-ki” applies to an NP that has taken a possessive marker or a locative case marker and produces an adjective. (This adjective is often immediately converted into a noun.) Therefore, it changes the grammatical category of its input, which is normally an indication of derivation. On the other hand, possessive markers and case markers clearly belong to the group of inflectional markers and ”-ki” comes after them. Since we would expect derivational markers to be closer to the stem, ”-ki” seems to also behave like an inflectional marker. Perhaps it does not fit into morphology at all, and it deserves to head its own construction: the ”ki-construction”. Until we can find more clues about this peculiar marker, we leave this issue aside. For the time being, suffice it to say that it does not matter much for us whether an affix is considered inflection or derivation, as long as we can account for its behavior with CCG.

(12) T/A/M markers

- (a) -di:  $(S \setminus NP_{\text{NOM}}) \setminus V: \lambda p \lambda a \lambda t \lambda w. (pat\ w) \wedge (t < t_0)$
- (b) -iyor:  $(S \setminus NP_{\text{NOM}}) \setminus V: \lambda p \lambda a \lambda t_1 \lambda t_2 \lambda w. (pat\ w) \wedge (t_1 < t) \wedge (t < t_2)$
- (c) -meli:  $(S \setminus NP_{\text{NOM}}) \setminus V: \lambda p \lambda a \lambda t \lambda w. \forall w_{mod} (congruent' speaker' w_{mod} t_0 w_0) \implies (pat\ w_{mod})$

(13) Agreement markers

- (a) -Ø:  $(S \setminus NP_{\text{NOM},3S}) \setminus (S \setminus NP_{\text{NOM}}): \lambda p \lambda a. pa$
- (b) -Ø:  $S \setminus (S \setminus NP_{\text{NOM}}): \lambda p. phe'$

Examples in (12) constitute just a small sample from a great variety of possibilities. T/A/M markers contribute to the semantics of a verb on several different dimensions. To reflect these contributions, we use  $t$  and  $w$  to denote time points and possible worlds, following Moens and Steedman (1988).

For now, our T/A/M markers convert V to S|NP, but after further analysis, we might prefer a different syntactic category. Again for now, agreement markers come in pairs; one just marks the person-number feature of the subject, while the other also fulfills the place of a dropped subject.

At least one T/A/M marker is obligatory in Turkish finite verbs. If we assume the existence of a third-person singular marker, agreement markers are also obligatory.

(14) Voice markers

- (a) -dir:  $(V \setminus NP) \setminus V: \lambda p \lambda x \lambda a \lambda t \lambda w. init' (p\ x\ t\ w) at\ w$

(b) -il:  $V \setminus (V \setminus NP)$ :  $\lambda p \lambda t \lambda w. p \text{anon}' t w$

Using voice markers as an example, we can demonstrate the use of Skolem terms. As cited in PFTL, Dowty (1981) makes an analysis of passivization, employing Skolem functions as "existentials taking narrow scope with respect to the operator". Also, the set of variables taken by the Skolem function may be empty, in which case the Skolem term becomes a Skolem constant, "behaving like a wide-scope existential". Our passive marker in (14) (b) makes use of a Skolem constant, *anon'* in just this way. This technique frees us from having to introduce quantifiers where they would only create notational clutter.

There is a strong case for voice markers' actually being derivational suffixes. First, they obviously change the argument structure, adding and removing the arguments of a verb. This means their application cannot wait until after some arguments are filled by NP's of the sentence; they must apply directly to the verb stem and determine the final argument structure. Second, they are closest to the verb stem and they come before all (other) verbal inflection. Moreover, they even come before the derivational marker "-mek". Perhaps the CCG representations in (14) should be updated with a double slash. In Section 5.5, we will return to the important distinction between inflection and derivation with respect to CCG application modes.

## 5 Preliminary Work

Our best hope is to come up with a mechanistic explanation of why derivational morphology is / must be restricted the way it is. Derivational processes seem to have special properties that distinguish them from other linguistic processes. These properties are most clearly reflected in their CCG representations. We believe that these are not just random realizations, but are actually the consequence of simple properties inherent in the concept of derivation. In this section, we first try to develop CCG entries for some selected derivational affixes and come up with preliminary generalizations. These generalizations will be the first step in our attempts to formally investigate the special properties of derivational processes. Second, we return to the claim that syntactic and morphological processes differ in the way they interact with the sentential structure. Finally, we look deeper into a particular inflectional process.

### 5.1 Deriving from Nouns

In Turkish, there are 31 derivational affixes that apply on nouns (Bozşahin (2018)). Only one affix expecting a noun host is marked as productive in the original source, JND\_SIZ, but we believe there are several others, if we slightly lower our standards of productiveness. Especially the JND class affixes seem to be more productive.

(15) NND.CI: Lexical entries for "muhassebe -ci -Ø gel -di -Ø"

- (a) muhassebe:  $N_{3S}$ : *accounting'*
- (b) -ci:  $N \setminus N$ :  $\lambda x \lambda a. deals\_with' x a$
- (c) -Ø:  $NP_{NOM} \setminus N$ :  $\lambda x. x$
- (d) gel:  $V$ :  $\lambda a \lambda t \lambda w. come' at w$
- (e) -di:  $(S \setminus NP_{NOM}) \setminus V$ :  $\lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (f) -Ø:  $(S \setminus NP_{NOM}) \setminus (S \setminus NP_{NOM})$ :  $\lambda p \lambda a. pa$

(16) VND.LAN: Lexical entries for "kız -Ø ev -len -di -Ø"

- (a) kız:  $N_{3S}$ : *girl'*
- (b) -Ø:  $NP_{NOM} \setminus N$ :  $\lambda x. x$
- (c) ev:  $N_{3S}$ : *home'*
- (d) len:  $V \setminus N$ :  $\lambda x \lambda a \lambda t \lambda w. acquire' x at w$
- (e) -di:  $(S \setminus NP_{NOM}) \setminus V$ :  $\lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (f) -Ø:  $(S \setminus NP_{NOM}) \setminus (S \setminus NP_{NOM})$ :  $\lambda p \lambda a. pa$

(17) JND.LI: Lexical entries for "tat -lı -y1 -ØØ iste -di -Ø"

- (a) tat:  $N_{3S}$ : *sweetness'*

- (b) -lı:  $(N/N) \setminus N: \lambda x \lambda y \lambda a. (have' x a) \wedge (y a)$
- (c) -yı:  $NP_{ACC} \setminus N: \lambda x. x$
- (d) -ØØ:  $NP \setminus NP: \lambda x \lambda a. (x a) \wedge (indefinite' a)$
- (e) iste:  $V \setminus NP_{ACC}: \lambda x \lambda a \lambda t \lambda w. demand' x a t w$
- (f) -di:  $(S \setminus NP_{NOM}) \setminus V: \lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (g) -Ø:  $S \setminus (S \setminus NP_{NOM}): \lambda p. phe'$

The marker -ØØ will be described in Subsection 5.6.

(18) AND\_CA: Lexical entries for ”dost -ça gülümse -di -Ø”

- (a) dost:  $N_{3S}: friend'$
- (b) -ça:  $(V/V) \setminus N: \lambda x \lambda p \lambda a \lambda t \lambda w. (be' x a t w) \wedge (pat w)$
- (c) gülümse:  $V: \lambda a \lambda t \lambda w. smile' a t w$
- (d) -di:  $(S \setminus NP_{NOM}) \setminus V: \lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (e) -Ø:  $S \setminus (S \setminus NP_{NOM}): \lambda p. phe'$

## 5.2 Deriving from Verbs

There are 55 derivational affixes that apply on verbs. 7 of these are marked as productive in Bozşahin (2018).

Most suffixes of the NVD class chooses a thematic role made available by the verb and converts the input to an entity that could assume that role. Others create the name of act.

(19) NVD\_ACAK (patient)

- (a) al-acak: The entity that is taken
- (b) ver-ecek: The entity that is given

(20) NVD\_AK (location)

- (a) dur-ak: The place where one stops
- (b) yat-ak: The place where one lies

(21) NVD\_AMAK (patient)

- (a) bas-amak: The entity that is stepped on
- (b) tut-amak: The entity that is grabbed

(22) NVD\_AN (agent)

- (a) bak-an: The one who looks
- (b) kap-an: The entity that catches

(23) NVD\_ANAK (agent for intransitives / theme for transitives)

- (a) gel-enek: The habits that have come
- (b) yet-enek: The skills that suffice
- (c) tut-anak: The entity that is held
- (d) öde-nek: The amount that is paid

(24) NVD\_CA (name of the act for intransitives / theme for transitives)

- (a) eğlen-ce: The act of having fun
- (b) dinlen-ce: The act of resting
- (c) sakın-ca: What is to be avoided
- (d) düşün-ce: What is thought

(25) NVD\_GAC (agent)

- (a) süz-geç: The entity that filters
- (b) büyüt-eç: The entity that magnifies

(26) NVD\_YIS (name of the act + a sense of manner)

- (a) yürü-yüş: The act of walking
- (b) uç-uş: The act of flying

The above set of examples demonstrate how NVD processes select the agent, patient, location, theme or the action itself as the content of the new noun. This selection is often made according to the arity of the host verb. The fact that this selection is quite consistent may give us important clues on how derivation interacts with the argument structure. After all, consistency in such processes usually stem from an underlying structural property.

We propose the following CCG representations as a first step.

(27) NVD\_ACAK: Lexical entries for "al -acak -lar -ı iste -di -Ø"

- (a) al:  $V \backslash NP_{ACC}: \lambda x \lambda a.give' x a t w$
- (b) -acak:  $N \backslash (V \backslash NP_{ACC}): \lambda p \lambda x \lambda t \lambda w.(p x a n o n' t w) \wedge (t > t_0)$
- (c) -lar:  $N \backslash N: \lambda x.multiple' x$
- (d) -ı:  $NP_{ACC} \backslash N: \lambda x.x$
- (e) iste:  $V \backslash NP_{ACC}: \lambda x \lambda a \lambda t \lambda w.demand' x a t w$
- (f) -di:  $(S \backslash NP_{NOM}) \backslash V: \lambda p \lambda a \lambda t \lambda w.(p a t w) \wedge (t < t_0)$
- (g) -Ø:  $S \backslash (S \backslash NP_{NOM}): \lambda p.phe'$

An important observation here is how the argument structure of a verb must be preserved even after a noun is derived from it. Once the argument structure of a verb is represented in a logical form, it is there to stay. We can fulfill the individual arguments, for instance, we may

reduce the number of bound variables by replacing them with Skolem terms. However, the argument structure is kept till the end, no matter what operations take place during the derivation process.

Grimshaw (1990) mentions the generally accepted dichotomy between result and process nominals, but suggests another dichotomy between complex event nominals and others. She claims that the real distinction between the two kinds of nominals comes from their having or lacking argument structures. She points out that if a nominal lacks aspectual analysis, it will also lack an argument structure; otherwise, it will have an argument structure.

This dichotomy has a clear parallel in CCG. When we derive a complex event nominal from a verb, it seems we keep the argument structure of the host verb. On the other hand, when the result is not a complex event nominal, we are not allowed to keep the argument structure. Perhaps this is an indication that complex event nominals are actually derived every time and never lexicalized, but other nominals are lexicalized and the argument structures from host verbs are bypassed. As an added benefit, nominals that are locally ambiguous in terms of denoting complex events are automatically disambiguated during parsing, because the existence of an argument forces the reading with the argument structure, and vice versa.

VVD suffixes seem to leave the argument structure unchanged, but only modify the semantic content. Their application depends on the verbal class in Moens and Steedman (1988) given in Table 3.

Class	Events		States
Consequence	Atomic	Extended	
+	Culmination: recognize, spot	Culm. Process: build a house	know, love
-	Point: hiccup, tap, wink	Process: run, swim, walk	

Table 3: Distribution of Derivational Affixes

Most of these suffixes convert an atomic event to a process, or at least contribute a sense of repetition and inconclusiveness of action.

(28) VVD\_AKLA

(a) dur-akla

(b) it-ekle

(29) VVD\_ALA

(a) eṣ-ele

(b) ṣaṣ-ala

(30) VVD\_IKLA



- (a) uyu-kla
- (b) dit-ikle

(31) VVD\_USTUR

- (a) it-iştir
- (b) ver-iştir

This change in sense has further consequences regarding inflection. Appropriateness of T/A/M markers for a verb is determined by the verbal class. As the verb moves to the process class, it is qualified for markers and adverbs that are compatible only with processes. When these markers and adverbs are used with some verbs, they do not sound weird, because the verb has assumed a secondary verbal class. Some verbs, on the other hand, do sound weird in a similar context. Due to the sense of inconclusiveness and continuity, the derived verbs also sound better with duration specifying adverbs.

(32) "uyu"

- (a) Çocuk az önce uyu-du: Culmination
- (b) Çocuk 3 saattir uyu-yor: Process
- (c) Çocuk her akşam erkenden uyu-yor: Habitual action
- (d) \*Çocuk az önce uyu-kla-di: Ungrammatical
- (e) Çocuk 3 saattir uyu-kl-uyor: Process
- (f) Çocuk her akşam

(33) "dur"

- (a) Polis az önce dur-du: Culmination
- (b) Polis 3 haftadır kapının önünde dur-uyor: State
- (c) Polis çocuğu görünce dur-akla-dı: Process
- (d) \*Polis 3 dakikadır dur-akl-ıyor: Does not work
- (e) Polis çocuğu her gördüğünde dur-akl-ıyor: Repeated / habitual action

Strangely enough, all VVD affixes except one, involve a CAUS concept according to Bozşahin (2018). We will test the validity of this claim and look for possible reasons. (34) shows a preliminary representation of a VVD affix.

(34) VVD\_AKLA: Lexical entries for "otobüs -Ø dur -akla -dı -Ø"

- (a) otobüs:  $N_{3S}$ : *bus'*
- (b) -Ø:  $NP_{NOM} \backslash N$ :  $\lambda x.x$
- (c) dur:  $V_{CULM}$ :  $\lambda a \lambda t \lambda w.stop' at w$

- (d) -akla:  $V_{PROC} \setminus V_{CULM}$ :  $\lambda p.p$
- (e) -di:  $(S \setminus NP_{NOM}) \setminus V$ :  $\lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (f) -Ø:  $(S \setminus NP_{NOM}) \setminus (S \setminus NP_{NOM})$ :  $\lambda p \lambda a.pa$

(35) demonstrates the representation of the JVD\_IK affix. This affix also contributes a sense of passivization, which is reflected on its logical form with the Skolem constant *anon'*.

(35) JVD\_IK: Lexical entries for "yık -ık duvar -ı gör -dü -Ø"

- (a) yık:  $V \setminus NP_{ACC}$ :  $\lambda x \lambda a \lambda t \lambda w. demolish' x at w$
- (b) -ık:  $(N/N) \setminus (V \setminus NP_{ACC})$ :  $\lambda p \lambda x \lambda a \lambda t \lambda w. (pa anon' t w) \wedge (t < t_0) \wedge (xa)$
- (c) duvar:  $N$ : *wall'*
- (d) -ı:  $NP_{ACC} \setminus N$ :  $\lambda x.x$
- (e) gör:  $V \setminus NP_{ACC}$ :  $\lambda x \lambda a \lambda t \lambda w. see' x at w$
- (f) -dü:  $(S \setminus NP_{NOM}) \setminus V$ :  $\lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (g) -Ø:  $S \setminus (S \setminus NP_{NOM})$ :  $\lambda p.phe'$

We present below, not one but two AVD affixes, due to their fascinating logical forms.

(36) AVD\_ARAK: Lexical entries for "gülümse -yerek gel -di -Ø"

- (a) gülümse:  $V$ :  $\lambda a \lambda t \lambda w. smile' a$
- (b) -yerek:  $(V/V) \setminus V$ :  $\lambda p_1 \lambda p_2 \lambda a \lambda t \lambda w. (p_1 at w) \wedge (p_2 at w)$
- (c) gel:  $V$ :  $\lambda a \lambda t \lambda w. come' at w$
- (d) -di:  $(S \setminus NP_{NOM}) \setminus V$ :  $\lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (e) -Ø:  $S \setminus (S \setminus NP_{NOM})$ :  $\lambda p.phe'$

(37) AVD\_IP: Lexical entries for "gülümse -yip git -ti -Ø"

- (a) gülümse:  $V$ :  $\lambda a \lambda t \lambda w. smile' at w$
- (b) -yip:  $(V/V) \setminus V$ :  $\lambda p_1 \lambda p_2 \lambda a \lambda t \lambda w \lambda t_1. (p_1 at_1 w) \wedge (p_2 at w) \wedge (t_1 < t)$
- (c) git:  $V$ :  $\lambda a \lambda t \lambda w. go' at w$
- (d) -ti:  $(S \setminus NP_{NOM}) \setminus V$ :  $\lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (e) -Ø:  $S \setminus (S \setminus NP_{NOM})$ :  $\lambda p.phe'$

### 5.3 Deriving from Adjectives

There are 17\* (possibly 18) derivational affixes that apply on adjectives. Adjectives act like predicates in their logical forms, by taking the leftmost place in a term. If we wanted to avoid such a representation, we could use some actual verb, such as "be", as the predicate and place the adjective between the verb and the noun, just like a verbal argument. The choice between these two alternatives does not seem to make much difference in terms of expressive power, so

we prefer the former alternative which is more convenient.

NJD\_Ø is not listed in Bozşahin (2018), but we present it here to give an idea how such an affix would work.

(38) NJD\_Ø: Lexical entries for "küçük -Ø -ü getir -di -Ø"

- (a) küçük:  $N/N: \lambda x \lambda a \lambda d \lambda l. (little' (at' d) a) (d > l) (xa)$
- (b) -Ø:  $N \setminus \setminus (N/N): \lambda p. pincontext'$
- (c) -ü:  $NP_{ACC} \setminus N: \lambda x.x$
- (d) getir:  $V \setminus NP_{ACC}: \lambda x \lambda a \lambda t \lambda w. bring' x at w$
- (e) -di:  $(S \setminus NP_{NOM}) \setminus V: \lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (f) -Ø:  $S \setminus (S \setminus NP_{NOM}): \lambda p. phe'$

At first sight, VJD class affixes can be divided into two groups. Four affixes in the first group are used to indicate increases in the level of gradable adjectives. (The fact that only increases are accounted for in this way may be worth an investigation on its own.) Two affixes in the second group indicate a change in the subject's view of the object. In other words, subject's view of the object changes towards a higher level on the scale indicated by a gradable adjective. The following is an example from the larger group of VJD affixes.

(39) VJD\_LAS: Lexical entries for "köpek -Ø güzel -leş -ti -Ø"

- (a) köpek:  $N: dog'$
- (b) -Ø:  $NP_{NOM} \setminus N: \lambda x.x$
- (c) güzel:  $N/N: \lambda x \lambda a \lambda d \lambda l. (pretty' (at' d) a) \wedge (d > l) \wedge (xa)$
- (d) -leş:  $V \setminus \setminus (N/N): \lambda p \lambda x \lambda t \lambda w. (px) \wedge (increase' dt w)$
- (e) -ti:  $(S \setminus NP_{NOM}) \setminus V: \lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (f) -Ø:  $(S \setminus NP_{NOM}) \setminus (S \setminus NP_{NOM}): \lambda p \lambda a. pa$

All JJD affixes except one modify the intensity of the adjective. Thus, they apply on gradable adjectives. Probably, they could all be marked +GRAD. (The exception is JJD\_MSAR that makes the derivations "iyi-mser", "kara-msar" and the like.) To reflect the change in the grade of the adjective, we introduce a second limit  $l_1$ . In the following example,  $l_1$  is a limit easier to satisfy than the original  $l$ . When JJD\_MSI is applied, the grade of the adjective attained by the noun  $d$  is not claimed to be over the original limit; it is now between a lower limit and the original limit:  $l_1 < d < l$ .

(40) JJD\_IMSI: Lexical entries for "mavi -msi renk -i gör -dü -Ø"

- (a) mavi:  $N/N: \lambda x \lambda a \lambda d \lambda l. (blue' (at' d) a) \wedge (d > l) \wedge (xa)$
- (b) -msi:  $(N/N) \setminus \setminus (N/N): \lambda p \lambda x \lambda a \lambda d \lambda l \lambda l_1. (px ad l_1) \wedge (d < l) \wedge (l_1 < l)$
- (c) renk:  $N: color'$

- (d) -i:  $NP_{ACC} \backslash N: \lambda x.x$
- (e) gör:  $V \backslash NP_{ACC}: \lambda x \lambda a \lambda t \lambda w. see' x at w$
- (f) -dü:  $(S \backslash NP_{NOM}) \backslash V: \lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (g) -Ø:  $S \backslash (S \backslash NP_{NOM}): \lambda p.phe'$

Only 3 AJD affixes exist. The following example is the more productive one, producing manner adverbials.

(41) AJD\_CA: Lexical entries for "çocuk -Ø hızlı -ca gel -di -Ø"

- (a) çocuk:  $N: kid'$
- (b) -Ø:  $NP_{NOM} \backslash N: \lambda x.x$
- (c) hızlı:  $N/N: \lambda x \lambda a \lambda d \lambda l. (quick' (at' d) a) \wedge (d > l) \wedge (xa)$
- (d) -ca:  $(V/V) \backslash (N/N): \lambda p \lambda v \lambda a \lambda t \lambda w. p(v at w)$
- (e) gel:  $V: \lambda a \lambda t \lambda w. come' at w$
- (f) -di:  $(S \backslash NP_{NOM}) \backslash V: \lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (g) -Ø:  $(S \backslash NP_{NOM}) \backslash (S \backslash NP_{NOM}): \lambda p \lambda a.pa$

#### 5.4 Deriving from Adverbs

There is only 1 affix that apply on adverbs, and that one affix is not productive either. It seems derivation from adverbs is almost not possible in Turkish. Possibly, deriving from Turkish adverbs is too complex or counter-intuitive for some reason and either it was never invented, or it gradually disappeared. Of course there might be other affixes that we have not noticed so far. If there is really a lack of adverb-based derivational morphology in Turkish, this gap might lead us to an insightful exploration.

Still we try to come up with the lexical entry for (demin) "-cek" below.  $t_d$  indicates a specific length of time shorter than which any duration is deemed short enough for invoking the concept "demin".  $p$  is as usual the placeholder for the predicate,  $x$  stands for the subject,  $t$  and  $w$  are for time and world, respectively.

(42) AAD\_CEK: Lexical entries for "demin -cek gel -di -Ø"

- (a) demin:  $V/V: \lambda p \lambda x \lambda t \lambda w \lambda t_d. (p x t w) \wedge (t < t_0) \wedge (t_0 - t < t_d)$
- (b) -cek:  $(V/V) \backslash (V \backslash V): \lambda p.p$
- (c) gel:  $V: \lambda a \lambda t \lambda w. come' at w$
- (d) -di:  $(S \backslash NP_{NOM}) \backslash V: \lambda p \lambda a \lambda t \lambda w. (pat w) \wedge (t < t_0)$
- (e) -Ø:  $S \backslash (S \backslash NP_{NOM}): \lambda p.phe'$

## 5.5 Revisiting Inflection vs. Derivation

We have already mentioned the claim in Sezer (1991) that syntactic processes operate on the thematic structure, while morphological processes operate on the argument structure. The arguments for a distinct a-structure given in Grimshaw (1990) were also briefly explained in Section 3.1. It is easy to see how inflectional processes are required by the syntax. On the other hand, derivational processes operate with fundamentally different principles. Our intuition is that the difference lies in their being responsive to different layers of linguistic structures. While inflection applies only after all arguments are fulfilled, derivation applies on the individual lexical item and changes its argument structure.

In fact, their behavior is a direct consequence of the different tasks assigned to each kind of morphological process. Inflection must ensure all constituents fit into a final syntactic structure, but it cannot do it before individual constituents are "completed", in other words, have all their arguments fulfilled. Derivation, on the other hand, must take place before arguments start to fall into place within a-structures, because it is tasked with changing those very a-structures. Derivation is capable of creating new lexical items; therefore, it must take precedence over all syntactic processes.

Again in Section 3.1, we have cited Sezer (1991) for the distinction between s-selection and c-selection. We believe this is a significant finding in explaining why derivational processes behave differently than inflectional processes. Inflectional processes, which uniformly apply or do not apply to an entire grammatical category, and which do not make much reference to the semantic contents of their hosts, seem to require only c-selection. Derivational processes, however, consider both the grammatical category and the semantic content of their hosts, thus employing s-selection.

These arguments have been implicit in the choices we have made in developing the CCG representations of selected Turkish affixes. Our choice between `\` and `\\` has never been arbitrary or inconsistent. We have always used `\` for inflectional processes, since they cannot be forced to satisfy the three conditions for `\\` application, as given in PFTL:

(43) PFTL assumptions on the choice of `\\`

- (a) Host is lexical.
- (b) Operation is application only.
- (c) Output is lexical.

We were also careful to consistently use `\\` for derivational processes, as we expected from them to always satisfy the conditions listed above. In fact, this set of conditions is a defining property of a derivational processes.

For instance, we must use single `\` for relativizer markers because they are syntactically autonomous. Not to prevent the possibility of unbounded relativization, we must avoid using `\\`

which would restrict the application too tightly. Also, as PFTL suggests, "The relative marker occupies the slot of the tense marker in the inflectional paradigm.". Indeed, they come after voice markers and prevent any T/A/M markers from application; which indicate their belonging to the class of inflectional processes. However, they also change the grammatical category of their host to adjective, which could signal a derivational process is underway.

Such controversies only put further emphasis on the necessity to be able to account for inflectional and derivational processes with similar mechanisms. They also demonstrate the adequacy of our approach and the toolkit provided by CCG towards the realm of morphology. At the end of the day, morphology indeed distributes itself over the grammar by employing \ and \\ as it sees fit (PFTL). But CCG accommodates all these possibilities without requiring special mechanisms.

## 5.6 A Deeper Glance at Inflection: Marking Indefiniteness in Turkish

While our main focus is on derivational morphology, some linguistic aspects that do not necessarily fall into this domain are bound to appear under the spotlight from time to time. One such issue has to do with the representation of indefiniteness in Turkish.

Some Turkish sentences exhibit a curious lack of accusative case on the main argument of their verb. The following examples are from Sezer (1991) pp.42-43:

(44) Accusative argument to a matrix verb

- (a) Ali kitabı okudu.
- (b) Ali kitap okudu.

Normally we would expect the argument "kitab" to be assigned accusative case. In fact, it must be assigned accusative case whether it is realized phonologically or not, due to the Case Filter (Cowper (1992)). So what makes this argument lose its phonological case marking? There has been a lot of effort directed towards explaining this phenomenon. For example, Sezer (1991) distinguishes abstract and morphological case in Turkish in order to account for similar observations. We think this approach weakens the theory by adding a possibly unnecessary degree of freedom to the rule set governing case marking.

At this point, we find it necessary to refer to established views on the issue. We are aware that the literature on Turkish definiteness overwhelmingly holds accusative case responsible for the overt marking of definiteness, and does not put forward any overt marking for indefiniteness. Referentiality and specificity are also denied any dedicated overt markers. We believe the established opinion may not be the best possible explanation for the phenomena related to Turkish definiteness.

Dede (1986) creates a list of six possibilities regarding definiteness and referentiality of Turkish NPs: {definite, indefinite, nondefinite} X {referential, nonreferential}. Then, based on

this list, he sets out to discover the strategies of Turkish speakers regarding this issue. At the very beginning, he claims that referentiality in Turkish is never overtly marked. On the other hand, definite direct objects are exceptionally marked with the accusative case ending. In fact, he claims, for a definite direct object, the accusative marker is mandatory.

Several points can be raised against these claims. First, why is definiteness only marked in accusative case; what about other cases? We need to explain why direct objects have a special status with respect to definiteness. Second, if it is only definiteness that is marked, why must the accusative case, which has an important role to play in the sentential structure, be omitted due to indefiniteness? Perhaps a more intuitive explanation would involve marking indefiniteness instead of definiteness.

Tura (1986) conducts a similar analysis on non-verbal sentences. She uses the same list of combinations regarding definiteness and referentiality of Turkish NPs. Although the chapter looks into these issues with a focus on discourse, it includes quite interesting and relevant claims from our point of view. One of these describes why the singular-plural distinction can be voluntarily skipped. In (45) (a), the speaker "is not interested in establishing a discourse referent" and he demonstrates this by "neutralizing the singular-plural distinction" and not using any determiners.

(45) Singular-plural distinction omitted due to non-referential use

- (a) Beşte otobüs vardı ama...
- (b) Otobüs beşteydi.

We believe the act of "neutralizing" is key here; the speaker is allowed to neutralize some distinctions when the context is appropriate. Here the singular-plural distinction is neutralized due to non-referentiality; perhaps case marking is also neutralized elsewhere, and who knows what else. We will soon return to this issue, but first let us go over some other familiar examples. Possessive markers also constitute an important and obvious example in the definiteness-indefiniteness dichotomy.

(46) Possessive marked dependent of a nominal

- (a) arının sokması
- (b) arı sokması

(47) Determiner in a possessive construction

- (a) kitabın kapağı
- (b) kitap kapağı

(48) Accusative argument to a nominal

- (a) kitabı okumak

(b) kitap okumak

In (46) (a), "arının" must denote a definite bee that have been recently observed by the speaker and the hearer; therefore, it is allowed to carry a possessive marker. On the other hand, (b) lacks the possessive marker and refers to an indefinite, unspecific bee. In the first examples of (47) and (48), we are talking about a specific book and its cover, while in (b), we are not.

It seems this phenomenon is not limited to accusative marked objects; possessive markers and other case markers are also prone to being removed by a similar operation.

(49) Dative argument to a matrix verb

- (a) bisiklete binmek
- (b) bisiklet binmek
- (c) antibiyotiğe başlamak
- (d) antibiyotik başlamak

(50) Ablative argument to a matrix verb

- (a) kaydırdıktan kaymak
- (b) kaydırak kaymak

"bisiklete" and "antibiyotiğe" in (49) (a) and (c) refer to any item or some definite item. We need more context to disambiguate their meaning and determine the level of definiteness they convey. However, examples (b) and (d) can only refer to an indefinite item. Similarly in (50), the use of "kaydırdıktan" in (a) may lend itself to a definite reading, but the lack of ablative marker in (b) prevents this possibility.

We have not been able to construct examples with locative or instrumental case assigned items.

We suspect there is a compulsory subtractive inflectional suffix marking indefiniteness. This suffix phonologically removes case / possessive markers. The noun phrase remains in the same case and takes its place as an argument in the sentence, but the case marking is not pronounced.

The operation is more clearly observed with possessives and accusatives, because these are arguments, and they have a closer relation with their governor. We feel there is a sense of definiteness attached to arguments. For adjuncts, it seems, the operation is not compulsory. This might be because a similar sense of definiteness is not automatically assigned to adjuncts.

Finally, case marking is necessary for making an item's role in the sentence explicit. As pointed out in Sezer (1991), items without case must always be to the left of their governor, as otherwise the lack of case marking creates ambiguity. The following set of examples is taken and extended from Sezer (1991) pp.42-43:



(51) Variable word order with an accusative argument to a matrix verb

- (a) Ali kitabı okudu.
- (b) Kitabı Ali okudu.
- (c) Ali okudu kitabı.
- (d) Ali kitap okudu.
- (e) \*Kitap Ali okudu.
- (f) \*Ali okudu kitap.
- (g) \*Ali çok ilginç bir kitabı okumuş.
- (h) Ali çok ilginç bir kitap okumuş.
- (i) \*Çok ilginç bir kitap Ali okumuş.
- (j) \*Ali okumuş çok ilginç bir kitap.

Example (a) is an ordinary SOV sentence with an accusative definite argument. Emphasis can be placed anywhere in the sentence. Examples (b) and (c) have accusative definite arguments and non-standard word orders, OSV and SVO respectively. In these two sentences, the emphasis is on S, which immediately precedes V. (d) contains an indefinite argument with the standard SOV word order. Due to the standard word order, "kitap" can easily be identified as the object, and there is no ambiguity. (e) and (f) following non-standard word orders (OSV and SVO, respectively) are ungrammatical, since the lack of case marking creates ambiguity. Example (g) demonstrates that the accusative marker should be omitted on indefinite objects. The determiner, "bir" in this example brings a sense of indefiniteness to "kitap". Comparing (g) with (h), it seems the accusative marker is incompatible with this sense of indefiniteness. The subtractive morphological operation that marks indefiniteness is perhaps obligatory. With the SOV word order in (h), the sentence is perfectly acceptable. Since the accusative marker is missing in (i) and (j), non-standard word orders lead to ambiguity in distinguishing between the subject and the object. This is why these two sentences are not acceptable.

Another good example to test the integrity of this hypothesis could be the following sentence and its like (Ozge (2019)):

(52) Accusative marked generic object

- (a) Çocuklar çikolatayı çok sever.

The noun phrase "çikolata-yı" obviously does not refer to any particular piece of chocolate, but could one say it is indefinite? At first sight, it seems to refer to the general concept of chocolate, or a generic sense of chocolate. In this sense, "chocolate" has just one instance and it is definite, thus maybe our hypothesis survives this example.

If we follow this path to the extreme, it can even be argued that plural markers are dropped due to indefiniteness. This way one could explain how a particular common noun can become a collective noun. Also, one could explain why a plural must either be definite or must cover

all instances of an item; if a plural is indefinite, it lacks the plural marker whether or not it is accompanied by a determiner.

(53) Plural nominative subject to a matrix verb

- (a) Japonlar yapmış.
- (b) Birkaç Japon yapmış.
- (c) Japon yapmış.

(54) Plural nominative subject to a matrix verb

- (a) Ambara koltuklar geldi.
- (b) Ambara birkaç koltuk geldi.
- (c) Ambara koltuk geldi.

(55) Plural accusative argument to a matrix verb

- (a) Simitleri aldım.
- (b) Birkaç simit aldım.
- (c) Simit aldım.

In (53) (a), "Japonlar", either denotes some known group of, or all instances of the class. In the former possibility, it is definite which group is indicated. In the latter possibility, since all instances of the class is referred to, we cannot speak of indefiniteness. "Koltuklar" in (b) only allows the former reading. Examples (b) in (53) and (54) contain subjects that are modified by "birkaç" indicating their indefiniteness and number. Despite their being plural, these nouns do not take plural markers. This can be explained by the obligatory indefiniteness marker that eliminates the plural marker. Examples (c) in these groups have subjects that are ambiguous between two readings: Either they denote a definite instance of a class (although OSV in (54) reduces the plausibility of this reading), or they denote an indefinite group of instances (somehow plurality is implied). The former case is business as usual. In the latter case, we can explain the lack of plural markers with subtractive morphology. The group of examples in (55) involve accusative markers in addition to the plural markers. In (a), we have a definite group of bagels, therefore both markers can be observed. In (b), indefiniteness is explicitly indicated by a determiner, thus both markers are dropped. Finally, "Simit" in (c) also lacks both markers, although indefiniteness is not made explicit by a determiner.

The third examples from each series can appear to be generic nouns, leading one to argue they should be exempt from an indefiniteness marker as in (52) (a). In (53) (c) "Japon" is after all an agent in an event, thus probably is a collection of people, rather than a generic concept. (54) (c) also cannot be a generic concept, because the physical arrival of a group of couches is described. (55) (c) is similar, but in this case, "Simit" is the object, rather than the subject. This final example also shows how the indefiniteness marker is capable of removing both the case marker and the plural marker in one strike. Perhaps the indefiniteness marker removes all

inflection.

Word order indeed seems to play a role in determining the level of definiteness. As the following examples demonstrate, a different word order may indicate a different level of definiteness.

(56) Plural nominative subject to a matrix verb

- (a) Koltuklar ambara geldi.
- (b) Birkaç koltuk ambara geldi.
- (c) Koltuk ambara geldi.

Compared to (54), we observe no meaning changes in (a) and (b) of (56). In (c), however, we previously observed two possible meanings in (54), while in (56) only one of them remains: It must be a single definite couch that arrived.

Applying the subtractive indefiniteness marker seems obligatory with some determiners indicating indefiniteness, but not always.

(57) The use with "bazı" determiner

- (a) ?Bazı Japonlar yapmış.
- (b) Ambara bazı koltuklar geldi.
- (c) İnternette bazı yorumlar gördüm.
- (d) Bazı takım elbiseli adamlar seni sordu.

(58) The use with "birkaç" determiner

- (a) \*Birkaç Japonlar yapmış.
- (b) \*Ambara birkaç koltuklar geldi.
- (c) \*İnternette birkaç yorumlar gördüm.
- (d) \*Birkaç takım elbiseli adamlar seni sordu.

It seems semantics plays an important role in determining the behavior of the indefiniteness marker. The determiner "birkaç" may be marking true indefiniteness, while "bazı" could signify indefiniteness in the speaker's context, but possible definiteness in the hearer's context. Instead, Bliss (2004) points out how "bazı" resembles "some" in English and always makes an NP indefinite. However, unlike "some", which may occur in context where the NP is specific or non-specific, "bazı" always makes the NP specific. Perhaps what we were looking for was hidden in the dichotomy between indefiniteness and non-specificity. There are not one, but two subtractive markers: the indefiniteness marker and the non-specificity marker. The former removes the case markers, while the latter removes the plural marker. The claim in Tura (1986) that the singular-plural distinction could be neutralized due to non-referentiality has a direct parallel in this solution. When both markers are present, we observe (55) (c) and the like. (Due to

limited time, we postpone making a deeper analysis in this direction to the next version of this text.)

Definiteness / Specificity	Specific	Non-specific
Definite	I'm looking for the manager, Ms Lee.	I'm looking for the manager, whoever that may be.
Indefinite	There's a certain word that I can never remember.	Think of a word, any word.

Table 4: Two dimensions of definiteness

Modified versions of example sentences in (59) provide evidence for this line of thinking. The "bazı" sentences sound somewhat more plausible combined with a question requesting information from the hearer. The accusative marker still drops due to indefiniteness, but the plural marker remains.

(59) The use with "bazı" determiner

- (a) Bazı Japonlar yapmış. Haberin var mı?
- (b) Ambara bazı koltuklar geldi. Haberin var mı?
- (c) İnternette bazı yorumlar gördüm. Haberin var mı?
- (d) Bazı takım elbiseli adamlar seni sordu. Haberin var mı?

A related issue concerns the subjects of relative clauses. Forms of object relative clause (ORC) subjects resemble genitive marked nouns. But are they really  $NP_{GEN}$ ? The following list is based on examples from PFTL.

(60) Relative clauses

- (a) Arının soktuğu kız
- (b) Arıların soktuğu kız
- (c) Bir arının soktuğu kız
- (d) \*Bir arıların soktuğu kız
- (e) Birkaç arının soktuğu kız
- (f) \*Birkaç arıların soktuğu kız
- (g) \*Bazı arının soktuğu kız
- (h) Bazı arıların soktuğu kız

(60) covers eight combinations on two dimensions. The first dimension has two levels controlling whether a plural marker is present. The second dimension has four levels: no determiner, "bir", "birkaç" and "bazı". There is ambiguity in (a) regarding definiteness, but it is somewhat closer to the reading that indicates a definite bee. In (b), there are multiple bees and they constitute a group identifiable by the speaker; perhaps they were flying around moments ago. In (c), however, the indefiniteness of "arının" is made explicit by the determiner "bir". Nevertheless,

the ”-ın” marker remains. (d) is obviously ungrammatical, due to the disagreement in number.

(e) is the next step, where ”birkaç” explicitly indicates there are multiple bees and they are indefinite. In this case, the plural marker drops, but the ”-ın” is still there. (f) is ungrammatical, because the sense of indefiniteness is present without a doubt, but the plural marker is used. We have previously demonstrated how the plural marker must drop in such cases, when the obligatory indefiniteness marker is applied. Finally in (g) and (h), we observe the opposite preference by the determiner; this time the plural marker must not be dropped, despite indefiniteness. As in our comments for (59), we cite Bliss (2004) and repeat the possibility of a second subtractive affix, that operates independently from the first one, and that represents a second dimension (specificity) in the space of indefiniteness. If we return to the main point, though, in both cases, the ”-ın” affix remains.

We can clearly observe that the ”-ın” marker should never be dropped, although in some cases indefiniteness is explicitly marked with ”bir”, ”birkaç” and ”bazı” determiners. If we assume our arguments about the subtractive indefiniteness marker are correct and ”-ın” is indeed the genitive marker, (60) would constitute counterexamples difficult to explain.

As cited in PFTL, George and Kornfilt (1981) explains why the ”-ın” in Turkish ORCs is not an instance of genitive case, but a marker of finiteness. Both these sources provide plenty of examples demonstrating the correctness of the claim. The following list presents our analyses of the underlying morphology for ORC examples. PL stands for the plural marker, INDEF for the indefiniteness marker and FIN for the finiteness marker. Finiteness marker naturally comes last, since it carries out its task high on the syntactic tree, linking the subject and the verb phrase. Parentheses indicate an affix removed by subtractive morphology.

(61) Relative clauses analyzed

- (a) Arının soktuğu kız: Arı-FIN ... / Arı-INDEF-FIN ...
- (b) Arıların soktuğu kız: Arı-PL-FIN ...
- (c) Bir arının soktuğu kız: Bir arı-INDEF-FIN ...
- (d) Birkaç arının soktuğu kız: Birkaç arı-(PL)-INDEF-FIN ...
- (e) Bazı arıların soktuğu kız: Bazı arı-PL-INDEF-FIN ...

In (61) (a) there is ambiguity in meaning. If we choose the reading with a definite bee, we must go with the first analysis, otherwise with the second one. In (e), the plural marker remains, despite being followed by the indefiniteness marker. The following set of examples demonstrate how the analyses would change if an unspecificity (UNSP) marker is also claimed. UNSP seems to come before INDEF. This time, it is UNSP that removes the plural marker, and not INDEF; therefore, (e) is not a problem anymore.

(62) Relative clauses analyzed

- (a) Arının soktuğu kız: Arı-FIN ... / Arı-UNSP-INDEF-FIN ...

- (b) Arıların soktuğu kız: Arı-PL-FIN ...
- (c) Bir arının soktuğu kız: Bir arı-UNSP-INDEF-FIN ...
- (d) Birkaç arının soktuğu kız: Birkaç arı-(PL)-UNSP-INDEF-FIN ...
- (e) Bazı arıların soktuğu kız: Bazı arı-PL-INDEF-FIN ...

If we are going to be strict and claim all kinds of inflection are always obligatory, we will need to mark specificity (SP) and definiteness (DEF), as well as unspecificity and indefiniteness. Also, there is no reason why a zero-marker (SG) should not be marking singular NPs. Perhaps what UNSP and INDEF do is simply removing any phonological realization of a certain class of affixes. For UNSP, this class is number markers, and for INDEF, case markers.

(63) Relative clauses analyzed

- (a) Arının soktuğu kız: Arı-SG-SP-DEF-FIN ... / Arı-SG-UNSP-INDEF-FIN ...
- (b) Arıların soktuğu kız: Arı-PL-SP-DEF-FIN ...
- (c) Bir arının soktuğu kız: Bir arı-(SG)-UNSP-INDEF-FIN ...
- (d) Birkaç arının soktuğu kız: Birkaç arı-(PL)-UNSP-INDEF-FIN ...
- (e) Bazı arıların soktuğu kız: Bazı arı-PL-SP-INDEF-FIN ...

If we make a preliminary attempt at constructing a sequence of NNI affixes, it would look like the following (We exclude ”-ki”, since it does not quite behave like an inflectional marker and rather seems to have a construction of its own.):

SG/PL - SP/UNSP - POSS - CASE - DEF/INDEF - FIN

## References

- M. Aronoff. *Morphology by itself: Stems and inflectional classes*. MIT press, 1994.
- M. Aronoff and K. Fudeman. *What is morphology?* Oxford: Blackwell Publishing, 2005.
- M. Baker. *Incorporation: A Theory of Grammatical Function Changing*. Chicago University Press, 1988.
- H. Bliss. The semantics of the bare noun in turkish. In I. Mezhevich and M. B. Dobrovolsky, editors, *Calgary Papers in Linguistics*, volume 25, pages 1–65. University of Calgary, 2004.
- C. Bozsahin. The combinatory morphemic lexicon. *Computational Linguistics*, 28(2):145–186, 2002.
- C. Bozsahin. *CCGlab Manual*, 2017. URL <http://bozsahin.github.io/ccglab>.
- C. Bozsahin. Some binary concepts in turkish morphology. Cem Bozsahin, 2018.
- A. Carstairs-McCarthy. *The evolution of morphology*. Oxford University Press, 2010.
- E. A. Cowper. *A concise introduction to syntactic theory : the government-binding approach*. University of Chicago Press, 1992.
- M. Dede. Definiteness and referentiality in turkish verbal sentences. In D. I. Slobin and K. Zimmer, editors, *Studies in Turkish Linguistics*, pages 147–164. John Benjamins, 1986. URL <https://www.jbe-platform.com/content/books/9789027279163-09ded>.
- D. Dowty. *The Scope of Lexical Rules*, chapter Quantification and the Lexicon, pages 79–106. Foris, Dordrecht, 1981.
- E. Erguvanli-Taylan, editor. *The Verb in Turkish*. Amsterdam: John Benjamins, 2001.
- L. George and J. Kornfilt. *Binding and Filtering*, chapter Finiteness and Boundedness in Turkish, pages 104–127. London: Croom Helm; Cambridge, Mass: MIT, 1981.
- A. Göksel and C. Kerslake. *Turkish : a comprehensive grammar*. London: Routledge, 2005.
- J. Grimshaw. *Argument Structure*. MIT Press, 1990.
- K. Hale and S. J. Keyser. *Prolegomenon to a Theory of Argument Structure*. MIT Press, 2002.
- C. Kennedy. Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30:1–45, 2007.
- C. Kennedy and L. McNally. Scale structure, degree modification, and the semantics of gradable predicates. *Linguistic Society of America*, 81(2):345–381, 2005.
- O. Kılıç and C. Bozsahin. Selection of linker type in emphatic reduplication: Speaker’s intuition meets corpus statistics. In *35th Annual Meeting of Cognitive Science Society*, volume 35, 2013.

- R. Lieber. *Deconstructing morphology: word formation in syntactic theory*. University of Chicago Press, 1992.
- M. Moens and M. Steedman. Temporal ontology and temporal reference. *Computational Linguistics*, 14:15–28, 1988.
- K. Oflazer, E. Göçmen, and C. Bozşahin. An outline of turkish morphology. Technical report, Middle East Technical University and Bilkent University, 1995. Re-issued in 2014.
- U. Ozge. Personal communication, 2019.
- F. Paoli. Comparative logic as an approach to comparison in natural language. *Journal of Semantics*, 16:67–96, 1999.
- C. Paradis. Adjectives and boundedness. *Cognitive Linguistics*, 12(1):47–65, 2001.
- F. E. Sezer. *Issues in Turkish Syntax*. PhD thesis, Harvard University, 1991.
- M. Steedman and J. Baldridge. *Non-Transformational Syntax*, chapter Combinatory Categorical Grammar, pages 181–224. Blackwell, 2011.
- M. Steedman and C. Bozşahin. Projecting from the lexicon. Mark Steedman and Cem Bozşahin, n.d.
- S. S. Tura. Definiteness and referentiality in turkish nonverbal sentences. In D. I. Slobin and K. Zimmer, editors, *Studies in Turkish Linguistics*, pages 165–194. John Benjamins, 1986. URL <https://www.jbe-platform.com/content/books/9789027279163-09ded>.
- R. D. Van-Valin-Jr. *Linguistic Universals*, chapter Some Universals of Verb Semantics, pages 155–178. Cambridge University Press, 2006.
- A. Wedel. Turkish emphatic reduplication. Technical report, UC Santa Cruz: Working Papers, 1999.