RELATIONAL-REALIZATIONAL SYNTAX: AN ARCHITECTURE FOR SPECIFYING AND LEARNING MORPHOSYNTACTIC DESCRIPTIONS

Reut Tsarfaty
Uppsala University

Proceedings of the LFG10 Conference

Miriam Butt and Tracy Holloway King (Editors)

2010

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

This paper presents a novel architecture for specifying rich morphosyntactic representations and learning the associated grammars from annotated data. The key idea underlying the architecture is the application of the traditional notion of a "paradigm" to the syntactic domain. N-place predicates associated with paradigm cells are viewed as relational networks that are realized recursively by combining and ordering cells from other paradigms. The complete morphosyntactic representation of a sentence is then viewed as a nested integrated structure interleaving function and form by means of realization rules. This architecture, called *Relational-Realizational*, has a simple instantiation as a generative probabilistic model of which parameters can be statistically learned from treebank data. An application of this model to Hebrew allows for accurate description of word-order and argument marking patterns familiar from Semitic traditional grammars. The associated treebank grammar can be used for statistical parsing and is shown to improve state-of-the-art parsing results for Hebrew. The availability of a simple, formal, robust, implementable and statistically interpretable working model opens new horizons in computational linguistics — at least in principle, we should now be able to quantify typological trends which have so far been stated informally or only tacitly reflected in corpus statistics.

1 Introduction

Precision grammars and treebank grammars present two alternatives for obtaining an accurate, consistent and maximally complete syntactic analysis of natural language sentences. Precision grammars are developed by trained linguists who seek to encode their observations and generalizations as formal statements and constraints, in order to objectively describe natural language phenomena and formulate predictions in terms of inductive hypotheses. Treebank grammars are developed in engineering-oriented natural language processing environments to satisfy the need of technological applications to access an abstract representation of sentences and pass it on to downstream modules for further (e.g., semantic) processing.

For a long time these two research endeavors have been conducted in separate communities and optimized for disparate goals. The development of precision grammars emphasizes stating explicitly how surface expressions map to abstract grammatical functions. A formal framework such as Lexical Functional Grammar (LFG), for instance, has been used to articulate theories on how surface forms are mapped to feature structures via an imperfect correspondence function in different languages (Bresnan, 2000). This is intended to reveal rules and regularities and to provide a realistic way to approach the study of linguistic *universals*. Treebank

[†]The research reported in this article has been conducted at the Institute for Logic, Language and Computation at the University of Amsterdam, and funded by the Dutch Science Foundation (NWO), grant number 017.001.271. The author is currently a researcher at Uppsala University, funded by the project 'Syntactic Analysis for Synthetic Languages'. The author is indebted to the participants of LFG 2010 for stimulating discussion and to Joakim Nivre for insightful comments on earlier drafts.

grammars (Charniak, 1996) are developed so that they are easy to acquire and robust in the face of noise. The development of treebank grammars is often sensitive to annotation idiosyncrasies and more often than not it does not reflect an articulated theory. The resulting analysis is useful in as much as it helps to recover predicate argument relations. From a downstream application point of view, no attention is required to the kind of formal means used to realize these relations.

Recently, these somewhat disparate research efforts started to acknowledge their usefulness for one another. On the one hand, augmenting simple treebank grammars with surfacy linguistic notions such as the position of the head (Collins, 2003) or topological fields (Kübler, 2008) was shown to improve the precision of these treebank grammars. On the other hand, acquiring deep, precision grammars that represent multiple layers of linguistic description from treebank data, e.g., by Hockenmaier and Steedman (2002), Miyao and Tsujii (2008), and Cahill et al. (2008), helped to improve the coverage of such deep grammars and increase their robustness in the face of noise. Notwithstanding, through these and other research efforts it has also become apparent that borrowing terms, constructs and techniques from one research vein and applying them to another may be too simplistic an approach for specifying and statistically learning complex linguistic phenomena.

A case in point is the use of *morphology* in parsing. Treebank grammars developed for configurational languages such as English do not always carry over to less configurational languages effectively because existing models assign probabilities based on configurational positions. Rich morphosyntactic interactions which are orthogonal to syntactic configurations do not get assigned their own probability mass. This entails that the competition between morphology and syntax cannot be fully materialized when such grammars are used, e.g., for parsing (Tsarfaty et al., 2010). On the other hand, a precision grammar such as LFG exploits a wide-range of dependencies in a parallel architecture which makes it challenging to assign a probabilistic interpretation to them. A probabilistic interpretation requires the explicit specification of correlations in the form of conditional independence. Because of the huge space of possible morphosyntactic combinations, trying to learn all the possible options for the integration of the different levels often leads to an explosion of the space of parameters, which in turns leads to extreme sparseness.

This paper takes a step back to consider the task from first principles and develops a novel architecture which remains faithful to both kinds of goals. The proposed architecture, called *Relational-Realizational*, is adequate for economically describing rich morphosyntactic interactions, and, at the same time, it can be used to define an interpretable grammar that can be read off of treebank data and may be used for efficient parsing. The key idea underlying the architecture is the proposal to apply the traditional notion of a “paradigm” to the syntactic domain. Syntactic constituents and their associated features are related to one another in the same way that the set of inflected word forms of a lexeme defines a paradigm. The feature combinations define different cells in the inflectional class of the paradigms, and the N-place predicate associated with paradigms are viewed as relational networks that are realized recursively by combining and ordering cells from other paradigms.

In this paper we show that viewing complex morphosyntactic representation of a sentence as a nested paradigmatic structure allows us to describe profound linguistic facts concerning Modern Hebrew morphosyntax and at the same time instantiates a generative probabilistic model that improves parsing results for Hebrew. The parameters that are learned from the data can be interpreted as capturing different dimensions of realization, and the parameter tables read off of the treebank can be shown to quantify observations about argument marking which have traditionally been stated qualitatively. While this modeling strategy shares a lot of underlying assumptions with LFG, its integrative and realizational nature opens new horizons in the attempt to marry the theoretical and the statistical approaches.

The remainder of this paper is organized as follows. Section 2 outlines the proposal to extend the paradigmatic view from the morphological to the syntactic domain, and shows how appropriate independence assumptions turn it into a generative probabilistic model that can be effectively used for statistical parsing. Section 3 demonstrates how this architecture can be used to describe argument marking patterns in Modern Hebrew. Here we focus on word-order and differential object-marking, but the same methodology carries over to other morphosyntactic phenomena such as agreement. Section 4 summarizes the results of a quantitative evaluation of the resulting treebank grammars and in section 5 we summarize the contribution and conclude.

2 The Model: Relational-Realizational

For a statistical model to meet the challenge of linguistic adequacy it is ultimately required to learn how abstract grammatical functions, such as *subject, object, past tense* or *grammatical gender*, are manifested through a range of language-specific forms, such as *word position*, *affixes*, *phrase-level agreement*, and so on. The view that *form* and *function* in natural language are independent from one another has been wide spread in typological studies (cf. Sapir (1921)), and is a fundamental principle motivating the projection architecture in LFG, where c-structures are mapped to f-structures through an *imperfect* correspondence function. On top of that, the idea that form and function are independent has been mastered and extensively utilized in theoretical morphology, where form-function separation guides the descriptions of the ways morphosyntactic representations map to words (Anderson (1992); Aronoff (1994); Blevins (2010); Matthews (1974); Stump (2001)).

Let us illustrate form-function independence in morphology. Consider, for instance, the realization of the grammatical property [+plural] in English. The property [+plural] in English is expressed in a variety of forms, such as ‘kids’, ‘children’, ‘men’, ‘sheep’, ‘oxen’, and so on. It falls out of this variation that the morphological exponent ‘s’ is not a necessary condition for the realization of [+plural]. At the same time, the exponent ‘s’ associated with English [+plural] expresses also the present-tense third-person singular property-bundle in the morphology of verbs (as in ‘eats’), so ‘s’ is not even sufficient for determining [+plural] in English.

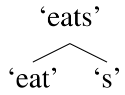


Figure 1: Morpheme-based morphology

Different morphological models approach differently the fundamental question how form-function association are stored, and how form-function relations between property-bundles and surface word forms are being computed. Stump (2001) singles out paradigmatic, realizational models as adequate for modeling complex form-function correspondence patterns in morphology. This paper articulates a proposal to extend this modeling strategy from the morphological to the syntactic domain. Here, instead of looking at grammatical *properties* such as *gender, number* etc., we consider grammatical *relations* such as *subject of*, *object of* and so on. Instead of analyzing property-bundles, we look at sets of relations that define complete argument structures. The result of this exercise leads to a paradigmatic and realizational model which is adequate for describing many-to-many, form-function correspondence patterns between sets of grammatical relations and surface syntactic structures which may be intertwined with complex morphology.

2.1 Models for Morphology

Morphologists seek to describe exponence relations — associations between properties of words and surface formatives — of several kinds. A **Simple Exponence** relation is a one-to-one relation between a property and a formative (such as [z] in 'seas' or [d] in 'sailed'). This clear association between formatives and properties is often the case in radically agglutinating languages, such as Turkish. **Cumulative Exponence** is a one-to-many relation common in fusional languages such as Latin, where a single ending may realize, e.g., number/case feature combinations. Many-to-one relations are called **Extended Exponence**, where the joint contribution of different formatives is required for expressing a single property. An example is the Greek verb *e-le-ly-k-e-te* where the perfective is marked by at least three forms: 'le', 'y' , and '-te' interleaved with other exponents (Matthews, 1974, p. 180).

In theoretical morphology, there exist at least two different ways to approach such descriptions. In the American structuralist tradition (Bloomfield (1933) and followers), a word form like 'eats' is seen as a combination of two different forms, a root 'eat' and a suffix 's'. These forms are defined to be *morphemes* — minimal meaningful units in the language linking sound to meaning in the Saussurian sense. The functional characterization of the word 'eats' is then derived from combining the functions of the parts, and its form is the result of their concatenation. This is the *Morpheme-Based (MB)* view of morphology, illustrated in figure 1.

A different way to analyze the word form 'eats' is to view the set {'eat', 'eats', 'ate', 'eaten'} as associated with a single lexical entry, a *lexeme* EAT. The word

/EAT/	1Sing	2Sing	3Sing	1Pl	2Pl	3Pl
Past	1SingPast	2SingPast	3SingPast	1PlPast	2PlPast	3PlPast
Present	1SingPres	2SingPres	3SingPres	1PlPres	2PlPres	3PlPres
Perfect	1SingPerf	2SingPerf	3SingPerf	1PlPerf	2PlPerf	3PlPerf

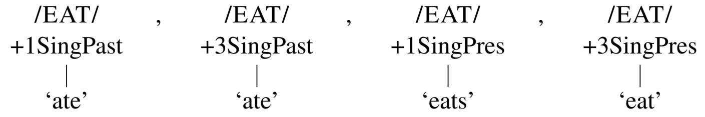


Figure 2: Word and Paradigm Morphology

forms {‘eat’, ‘eats’, ‘ate’, ‘eaten’} then realize different abstract descriptions associated with the lexeme; ‘the past form of EAT’, ‘the present tense third-person singular form of EAT’, etc. Under this view, this set is a concrete realization of a *paradigm* that has cells that associate a lexeme with different combinations of morphosyntactic representations of the feature values of *tense, gender* and *number*. This is the *Word-Based (WB)* view of morphology.

The graphical depiction of this alternative strategy is shown in figure 2. The paradigm defines different content cells, i.e., the different functions that a word-form associated with this lexeme may fill up in the syntactic structure. The mechanism used to manipulate the set of properties (e.g., syntactic requirements) is assumed to be distinct of the mechanisms that construct forms (e.g., a finite-state machine). *Morphological paradigms* define how the different word-forms are related, and *realization rules* define the mapping between well-formed property-bundles defining paradigm cells and the appropriate word-forms that make up the paradigm.

Stump (2001) isolates two modeling assumptions that jointly characterize the differences between these two opposing views. Distinguishing *lexical* and *inferential* models is concerned with how the association of properties to exponents is stored in the grammar, and the distinction between *incremental* and *realizational* approaches is about the ways multiple properties get associated with word-forms.

- ***Lexical* vs. *Inferential* Approaches** In *lexical* approaches, form is primary to function. Forms are listed in a lexicon where they are associated directly with discrete functions. In *inferential* approaches, functions are primary, and the model explicitly computes the associated forms in the course of analysis.

- ***Incremental* vs. *Realizational* Approaches** In *incremental* models, properties are accumulated incrementally. In *realizational* ones, complete property-bundles are the precondition for, rather than the outcome of, the application of spell-out rules. In *incremental* models, words are artifacts. In *realizational* ones, words have an independent formal status beyond the combination of parts, and they are related to one another through the notion of a paradigm.

Morpheme-based (MB) approaches, as alluded to above, are *lexical* and *incremental.* They constitute the simplest, most intuitive way to view morphology. Describing morphological patterns involving formatives that go beyond simple exponence, however, may become cumbersome with MB approaches. An example comes from null realization. In the case of “sheep”, for instance, it is necessary for *lexical-incremental* approaches to stipulate an empty formative and associate it with a [+plural] property. Empirical evidence for such ‘empty morphemes’ is hard to establish, but without it the description collapses. Another challenge faced by *lexical-incremental* models is the ‘selection problem’, that is, the challenge of incrementally choosing morphemes that ‘go together’ when generating feature combinations in the lexicon. In actuality, feature occurrences may be interdependent.

Old prescriptive grammars never face such challenges because they invoke *Word and Paradigm (W&P)* approaches which are *inferential* and *realizational*. Their modern conception, *Extended Word and Paradigm (EW&P)* approaches (Anderson, 1992; Stump, 2001) define a paradigm by means of an abstract lexeme and a set of well-formed feature-bundles a priori associated with the lexeme. The morphosyntactic representation of cells in the paradigm is primary to word forms. Well-formed property-bundles representing these cells are delivered to the morphological model (say, by the syntax), and the morphological component consists of a set of realization rules which *interpret* these property-bundles. This interpretation may specify nothing (as is the case in ‘sheep’ plural) or it may articulate mapping of a subset of the property to noncontiguous parts of a form, as in the Greek case.

Lexical-incremental theories then work to perfection in cases of radical agglutination, but face challenges with more complex morphology. The flexibility of combining paradigmatic associations with realization rules in *inferential-realizational* approaches, however, makes them well suited for describing exponence relations of all kinds, as seen in Anderson (1992), Aronoff (1994), Blevins (2010) and more.

2.2 Models for Syntax

It should be clear from the outset that syntactic analysis of natural language sentences needs to cope with a range of exponence relations which is as diverse as morphological exponence. By Syntactic Exponence I refer to the relationship between abstract grammatical relations and their surface manifestation. **Simple Exponence** is a one-to-one relation between abstract entities and configurations, such as the relation between an NP dominated by an S and the *subject* function, as is articulated in early versions of X-bar theory. **Cumulative Exponence** is the realization of multiple syntactic functions by means of a single syntactic formative; this happens, for instance, in structures involving *clitics*, phonologically reduced elements that indicate an independent grammatical entity in addition to the one associated with their host. **Extended Exponence** is a many-to-one relation between formatives and functions, manifested through, e.g., periphrasis, functional co-heads (such as AUX in Warlpiri) or referring to the morphology of multiple forms, e.g., in differential marking or agreement. Given the diversity of exponence relations it may be fruitful to identify general principles according to which one could derive modeling strategies to describe syntactic exponence relations. This paper asks two orthogonal questions, parallel to the ones that concerned models for morphology: firstly, how the model stores form-function associations, and secondly, how complete sets of relations and properties get associated with the morphosyntactic structures that realize them.

- ***Configurational* vs. *Relational* Approaches**
  In *configurational* approaches configurations are primary to functions. Configurational pieces are used to define grammatical functions, and grammatical relations are derived notions. *Relational* approaches take grammatical relations as primary and primitive and separate them from their surface manifestation. The model then calculates form in the course of the analysis.

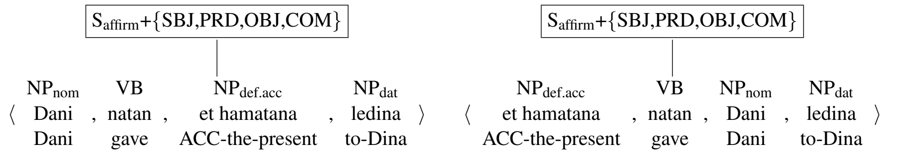
- ***Incremental* vs. *Realizational* Approaches**
  In *incremental* approaches, the theoretical primitives (configurational or relational) are accumulated incrementally in the course of the syntactic analysis. Argument-structure is an artifact of the combination of syntactic pieces. In *realizational* approaches, complete sets of primitives are a precondition for, rather than the outcome of, the application of syntactic realization rules. Argument-structure then has a formal status beyond the sum of its parts.

The *configurational-incremental* view is compatible with configurational languages, where associations of configurational positions and grammatical relations are very tight. In nonconfigurational languages, configurational positions do not always stand in one-to-one correspondence to grammatical relations, and morphological exponents in the syntactic structure may alter or supplement form-function associations. To effectively capture argument marking patterns that have to do with the interactions of configurations with complex morphology, we propose to extend the architectural design of the W&P approach to the syntactic domain. The model is thus required to be (i) *relational*, i.e., function is primary to form, grammatical relations are primary to complete surface structures, and (ii) *relatizational*, i.e., a complete set of functions is a pre-condition for the application of realization rules which interpret them as morphosyntactic forms. The key idea underlying the proposal is that forms of higher-level constituents in a syntactic structure define the function of lower-level constituents, and the recursive realization process goes on to unfold the surface structure. As the structure unfolds, feature-bundles become increasingly specific until they can be fed into a model of W&P morphology.

2.3 Relational-Realizational Syntax

The design of the relational-realizational architecture starts off from an assumption that has been pertinent in syntactic theory since Relational Grammars (Postal and Perlmutter, 1977), and which has also inspired the architectural design of LFG — that grammatical relations are theoretical primitives, and that the description of

S⟨PRED⟩ FEATS ARG-ST	Affirmative	Interrogative	Imperative
intransitive	S_{affirm}+{SBJ,PRD}	S_{inter}+{SBJ,PRD}	S_{imper}+{SBJ,PRD}
transitive	S_{affirm}+{SBJ,PRD,OBJ}	S_{inter}+{SBJ,PRD,OBJ}	S_{imper}+{SBJ,PRD,OBJ}
ditransitive	$\boxed{S_{affirm}+\{SBJ,PRD,OBJ,COM\}}$	S_{inter}+{SBJ,PRD,OBJ,COM}	S_{imper}+{SBJ,PRD,OBJ,COM}



Saffirm+{SBJ,PRD,OBJ,COM}
NPnom VB NPdef.acc NPdat
⟨ Dani , natan , et hamatana , ledina ⟩
Dani gave ACC-the-present to-Dina
Saffirm+{SBJ,PRD,OBJ,COM}
NPdef.acc VB NPnom NPdat
⟨ et hamatana , natan , Dani , ledina ⟩
ACC-the-present gave Dani to-Dina

Figure 3: The Relational-Realizational Architecture

how they are realized can vary from one language to another. This is the *relational* assumption. In RGs, Relational Networks (in LFG, N-place predicates) group together sets of relations which are associated with the complete clauses or sentences that realize them. This is the *realizational* assumption. Here we aim to define a generative device that allows us to generate such function-to-form associations in an integrated representation. This generative-integrated view of syntax will then be instrumental in turning the formal description into a proper probabilistic model.

2.3.1 Paradigmatic Organization

The intuitive idea of viewing syntax in terms of paradigms goes a long way back. Pike (1963) has shown that describing the syntactic constructions in a language by means of a feature-based paradigm, as was used for phonological descriptions at that time, provides for an economic and intuitive way to compare the grammar of different languages. Matthews (1981) suggested the notion of a paradigm to capture a set of syntactic alternations that are transformationally related. Here we extend this intuitive idea, that the category S represents a syntactic paradigm, to all syntactic categories. We abstract away from transformations as a mechanism for realizing the paradigmatic alternations, later to be replaced by *realization rules*.

We propose to associate syntactic categories with a feature-geometry that defines the functions that phrases of different types may fill. A syntactic inflectional class associated with a lexical predicate (PRED) and a feature geometry defines a syntactic paradigm. The lexical predicate designates a set of grammatical relations, here conceived as the *initial* level of relational networks (RNs) in the sense of RG. The features defined by the inflectional class may refine the network of arguments to be realized, here conceived as the *final* level of an RN in RG. Syntactic constituents are instantiations of particular cells in syntactic paradigms, each of which realizes a set of abstract relations and properties that defines the function of this constituent. Figure 3 illusrates a paradigm associated with constituents of type S.

2.3.2 Realization Rules

A syntactic category, a property-bundle and a lexical predicate designate a cell in a syntactic paradigm. For each cell in the paradigm, we would like to specify how this overall function is realized in the syntactic structure. Focusing on the realm of morphosyntax,[1] there are at least three ways in which a grammatical function may be locally realized: by designating certain positions for the realization of a relation (e.g., in SVO languages), by delegating a property to a dominated constituent (e.g., by simple case marking), and/or by distributing properties to a set of dominated constituents that stand in a certain relation (e.g., agreement).

Realization rules are formal rules that map the abstract representation of constituents as cells in paradigms onto a sequence of partially specified cells in other paradigms. Figure 3 shows an exemplar S paradigm and realization rules that implement two possible ways in which a particular content cell in the paradigm may be realized. The sequence of labels and features specify regions in other syntactic paradigms, which in turn may be associated with their own lexical predicate and relational network, and this realization process proceeds until the paradigm cells are fully specified and may be handed over to a model of W&P morphology.

While the organization principles of categories as paradigms is assumed to be universal, the ways in which the high-level category can group, order or distribute features to other paradigm cells are language-specific. It is a property of the language, rather than a property of the formal architecture, at which level of the hierarchy (clause, phrase, word) complete morphosyntactic representations (henceforth, MSRs) are handed over to morphology. This modeling strategy maintains a unified view of morphology and syntax that cuts across the separation between form and function, and draws the distinction between them according to the nature of the realization rules that spell out the function: syntax involves *recursion* to other paradigms, morphology maps functions directly to surface forms in the lexicon.

2.3.3 Independence Assumptions

There are different ways to specify realization rules that relate content cells in paradigms to sequences of content cells of dominated forms. This work proposes an approach prominent in morphology – to identify abstract units that can be combined to form new hypotheses and retain generalizations about how surface forms come about. These units need not be minimal Saussarian signs, but rather, they are different aspects of the complete form-function association. In order to identify these different aspects we isolate different dimensions of description for each local constituent, and point to independence assumptions between these dimensions. Such independence assumptions will lead straight forwardly to a probabilistic interpretation of the generative model. We articulate two independence assumptions for each paradigm cell: (i) the independence of form and function, and (ii) independence of different dimensions of realization.

[1] For the time being, we are discarding other means of realization such as intonation, prosody, etc.

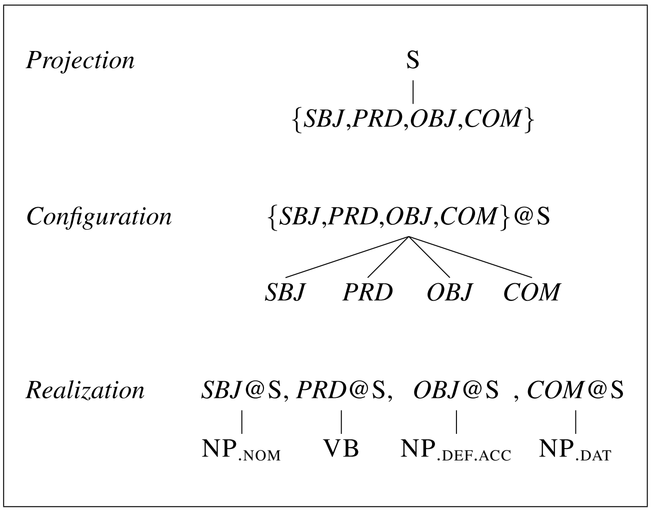


Figure 4: The Relational-Realizational (RR) backbone

These independence assumptions give rise to three generative phases for each rule, termed *projection, configuration* and *realization*, illustrated in figure 4.

The Projection Phase. The goal of the first phase in the realizational cycle is to pick out a content cell in the paradigm that specifies the function of a constituent. Let us assume a syntactic paradigm associated with a lexical predicate, a relational network, and a property-bundle. In the projection phase, a morphosyntactic representation of the syntactic category and the lexical predicate projects the set of grammatical relations that represents the set of arguments to be realized.

The Configuration Phase. Having picked out a cell in a syntactic paradigm, the remaining challenge is to spell out how it is realized. The configuration phase determines the ordering and juxtaposition of *relational slots*; these are slots in which different grammatical relations are realized. The configuration phase is at the same level of abstraction as the basic definition of word-order parameters in Greenberg (1963) – the order defines functions, not types of constituents. Furthermore, the realization of grammatical relations may be supplemented by additional elements such as auxiliaries or function words, co-heads, punctuation marks, and even slots reserved for modification or adjunction. So the configuration phase may reserve additional realizational slots spread out in between the relational labels.

The Realization Phase. The configuration phase allocated slots in which grammatical relations are to be realized as lower-level constituents. In order to realize these relations, we need to specify the syntactic category of the dominated constituents, and the features which are required to be marked within the scope of these constituents for realizing their function in the overall structure. Each of the relational and realizational slots is assigned a complete morphosyntactic representation that isolates a region in another syntactic paradigm.

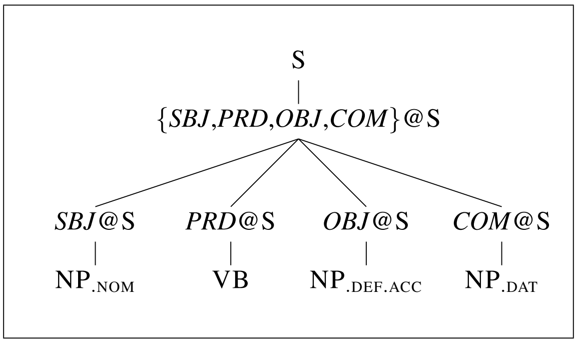


Figure 5: The Relational-Realizational (RR) representation

The properties delegated to the morphosyntactic representation may later be realized periphrastically (as a part of the configuration of a lower-level constituent) or morphologically (by delegating features to the morphosyntactic description of dominated constituents). As the structure unfolds, the morphosyntactic representation of constituents becomes increasingly specific until fully-specified descriptions may be handed over to the morphological component of the grammar for spell-out.

The Formal Model The resulting representational format of these three phases is a *linearly-ordered labeled tree* as depicted in figure 5. The complex labels of non-terminal nodes represent three distinct kinds of concepts: (i) grammatical relation labels (GRs) (ii) sets of grammatical relations, and (iii) sequences of morphosyntactic representations (MSRs) of constituents (marked here in indexed upper-case letters designating P(arent) and C(hildren)). We can identify in such trees context-free rules that correspond to the *projection*, *configuration* and *realization* phases which jointly spell-out syntactic constituents. The result of this process delivers MSRs to the next level of constituents. Morphological spell-out finally maps MSRs of pre-terminal constituents to surface forms.

(1)
- **Projection**
  $P \rightarrow \{GR_i\}_{i=1}^{n}$
- **Configuration**
  $\{GR_i\}_{i=1}^{n}@P \rightarrow \langle ..[GR_{i-1} : GR_i], GR_i, [GR_i : GR_{i+1}].. \rangle$
- **Realization**
  - for Relational Slots
    $\{GR_i@P \rightarrow C_i\}_{i=1}^{n}$
  - for Realizational Slots
    $\{GR_{i-1} : GR_i@P \rightarrow C_{i_{1_i}} \ldots C_{i_{m_i}}\}_{i=1}^{n+1}$
- **Spell-Out**
  $C \rightarrow s$

The Probabilistic Model Assuming the formal generative device just described, we can define probability distributions by choosing dependencies between phases of generation. Here each phase is conditionally dependent on the previous one.

(2)
- The **Projection** Distribution
  $\mathbf{P}_{\text{projection}}(\{GR_i\}_{i=1}^{n} \mid P)$
- The **Configuration** Distribution
  $\mathbf{P}_{\text{configuration}}(\langle..[GR_{i-1}{:}GR_i], GR_i, [GR_i{:}GR_{i+1}]..\rangle \mid \{GR_i\}_{i=1}^{n}@P)$
- The **Realization** Distribution
  - for Relational Slots
    $\mathbf{P}_{\text{realization}}(C_i \mid GR_i@P)$
  - for Realizational Slots
    $\mathbf{P}_{\text{realization}}(C_{i_{1_i}} \ldots C_{i_{m_i}} \mid GR_{i-1}{:}GR_i@P)$
- The **Spell-out** Distribution
  $\mathbf{P}_{\text{spellout}}(s \mid C)$

Because of the local independence between these parameters, the probability distributions can be estimated using relative frequency estimates that maximize the likelihood of the data. The estimated RR-PCFG can be used for efficient exhaustive parsing. Because conditional dependencies allow us to specify systematic relations between form and function, the model can also be used for describing consistent and complete argument marking patterns, as we do next.

3 The Application: Modern Hebrew Morphosyntax

The Relational-Realizational (RR) architecture defined in section 2 can be straightforwardly applied to describing morphosyntactic phenomena in the Semitic language Modern Hebrew. We show that the RR representation can capture linguistic facts about word order and argument marking in Hebrew, and that individual parameters can be used to state linguistic generalizations in a probabilistic grammar.

Word-Order and Sentence Structure Modern Hebrew is an SVO language, like English and many other languages. Its unmarked, canonical word-order pattern is *subject, verb, object* as in example (3).

(3) a. דני נתן את המתנה לדינה.
dani natan et hamatana ledina.
Dani gave ACC the-present to-Dina.
“Dani gave the present to Dina.”

On top of this basic word-order pattern, grammatical elements may be fronted, triggering an inversion of the unmarked Subject-Verb order (called *triggered inversion (TI)* in Shlonsky and Doron (1991)) as in (4a)-(4b). TI is similar to V2 constructions in Germanic languages. A *triggered inversion* stands in contrast with *free inversion*, in which subject-verb inversion may occur independently of such fronting (Shlonsky and Doron, 1991, footnote 2). Under certain information structuring conditions, *verb-initial* sentences are also allowed (VI in Melnik (2002)). A variation in the basic word order may also occur due to, e.g., *topicalization*, in which an element is fronted without triggering Subject-Verb inversion, as in (4c).

(4) a. את המתנה נתן דני לדינה.

et hamatana natan dani ledina.
ACC the-present gave Dani to-Dina.

"Dani gave the present to Dina."

b. לדינה נתן דני את המתנה.

ledina natan dani et hamatana.
to-dina gave Dani ACC the-present.

"Dani gave the present to Dina."

c. את המתנה, דני נתן לדינה.

et hamatana, dani natan ledina.
ACC the-present, Dani gave to-Dina.

"Dani gave the present to Dina."

The four alternative sentences in (3)–(4) only vary in their word-order pattern, due to triggering, inversion, and topicalization. The left hand side of figure 6 presents the RR representation of the structure of the different alternatives, and the right hand side shows the decomposition of the structure into generative phases, described as *parameters*. All sentences have type identical *projection* parameters and *realization* parameters, capturing the fact that their argument structure and argument marking patterns are exactly the same. The different sentences vary in the *configuration* parameters, reflecting the flexibility in the word-order pattern and additional realizational slots (e.g., punctuation). Learning these parameters from data can quantify exactly the production probabilities of the realization alternatives.

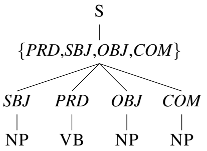
Differential Object Marking Core case marking in Hebrew displays sensitivity to the semantic properties of the phrase. This is a *differential* pattern of marking — objects in Hebrew are marked for accusativity if and only if they are also definite (Aissen, 2003). This pattern of marking is independent of the configurational positions of the different elements, as shown in (5).

(5) a. דני נתן את המתנה לדינה.

dani natan **et** **ha**matana ledina.
Dani gave **ACC** **DEF-**present DAT-Dina.

"Dani gave the present to Dina."

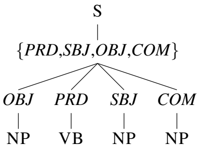
(3)



⇒

$$\mathbf{P}_{\text{projection}}(\{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\} \mid \mathrm{S})$$
$$\boxed{(\mathbf{P}_{\text{configuration}}\langle \mathit{SBJ}, \mathit{PRD}, \mathit{OBJ}, \mathit{COM}\rangle \mid \{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\}@\mathrm{S})}$$
$$\mathbf{P}_{\text{realization}}(\mathrm{VB} \mid \mathit{PRD}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{SBJ}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{OBJ}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{COM}@\mathrm{S})$$

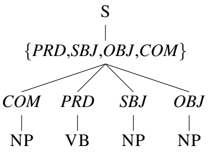
(4a)



⇒

$$\mathbf{P}_{\text{projection}}(\{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\} \mid \mathrm{S})$$
$$\boxed{\mathbf{P}_{\text{configuration}}(\langle \mathit{OBJ}, \mathit{PRD}, \mathit{SBJ}, \mathit{COM}\rangle \mid \{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\}@\mathrm{S})}$$
$$\mathbf{P}_{\text{realization}}(\mathrm{VB} \mid \mathit{PRD}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{SBJ}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{OBJ}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{COM}@\mathrm{S})$$

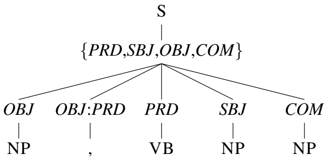
(4b)



⇒

$$\mathbf{P}_{\text{projection}}(\{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\} \mid \mathrm{S})$$
$$\boxed{\mathbf{P}_{\text{configuration}}(\langle \mathit{COM}, \mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}\rangle \mid \{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\}@\mathrm{S})}$$
$$\mathbf{P}_{\text{realization}}(\mathrm{VB} \mid \mathit{PRD}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{SBJ}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{OBJ}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{COM}@\mathrm{S})$$

(4c)



⇒

$$\mathbf{P}_{\text{projection}}(\{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\} \mid \mathrm{S})$$
$$\boxed{\mathbf{P}_{\text{configuration}}(\langle \mathit{OBJ}, \mathit{OBJ} : \mathit{PRD}, \mathit{PRD}, \mathit{SBJ}, \mathit{COM}\rangle \mid \{\mathit{PRD}, .., \mathit{PRD}\}@\mathrm{S})}$$
$$\mathbf{P}_{\text{realization}}(\mathrm{VB} \mid \mathit{PRD}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{SBJ}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{OBJ}@\mathrm{S})$$
$$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{COM}@\mathrm{S})$$
$$\boxed{\mathbf{P}_{\text{realization}}(, \mid \mathit{OBJ} : \mathit{PRD}@\mathrm{S})}$$

Figure 6: Basic word order and sentence structure

b. את המתנה נתן דני לדינה.
et **ha**matana natan dani ledina.
ACC **DEF-present** gave Dani DAT-Dina.
"Dani gave the present to Dina."

Via a particular sort of Semitic construction, termed Construct State Noun (CSN), an object phrase may become arbitrarily long. Such CSNs are also subject to differential marking patterns, however accustivity is marked at the beginning of the CSN and definiteness is marked on the last form. This means that there can be an unbounded distance between these inter-dependent feature markers, which is again orthogonal to the configurational position of the object, as shown in (6).

(6) a. דני נתן את מתנת יום ההולדת לדינה.
dani natan [**et** matnat yom **ha**huledet] ledina.
Dani gave [**ACC** present-of day-of **DEF**-birth] to-Dina.
"Dani gave the birthday present to Dina."

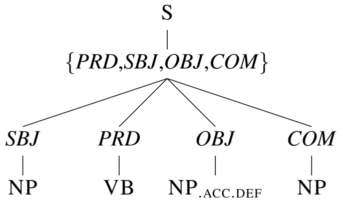
b. את מתנת יום ההולדת נתן דני לדינה.
[**et** matnat yom **ha**huledet] natan dani ledina.
[**ACC** present-of day-of **DEF**-birth] gave Dani to-Dina.
"Dani gave the birthday present to Dina."

The empirical facts are then as follows. Object marking in Hebrew requires reference to two overt markers, accusativity and definiteness. The contribution of the different markers is not independent, even though they appear on surface forms that are disjoint. This pattern of marking is orthogonal to the object position as well as to the way the object is spelled out (as a pronoun, noun, a CSN, etc.).

Let us consider sentences (5a)–(5b). The RR representation and parametrization of these constituents are presented in figure 7. Again, the difference between the parameter types lies at the configuration layer, but here we focus on the similarities. The two sentences share the *OBJ* relation parameter, $\mathbf{P}_{\text{realization}}(\text{NP} \mid \textit{OBJ}@\text{S})$. The label NP refers to an entire paradigm, but instead of NP we wish to indicate a morphosyntactic representation that isolates the functionally relevant region in the NP paradigm for an *OBJ* relation, so we specify $\mathbf{P}_{\text{realization}}(\text{NP}_{\text{DEF.ACC}} \mid \textit{OBJ}@\text{S})$.

The $\text{NP}_{\text{.DEF.ACC}}$ region then poses refined morphosyntactic requirements for this dominated constituent, regardless of its position. There are different ways in which these requirements can be filled. The $\text{NP}_{\text{.DEF.ACC}}$ MSR may be spelled out synthetically, for instance, using a pronoun marked for accusativity, gender, person, number and inherently definite; or it can be spelled out periphrastically, using the special accusative clitic את ('et') and a common noun marked for definiteness. It can also be spelled out syntactically, where the special clitic את attaches to an NP that has its own network of relations, e.g.,genitive constructions and CSNs, where the latter case is subject to a distinct feature-spreading pattern. In all cases, the isolated region in the NP paradigm makes sure that the realization is consistent and complete with respect to the delegated function.

(5a)



⇒ $\mathbf{P}_{\text{projection}}(\{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\} \mid \mathrm{S})$

$\boxed{\mathbf{P}_{\text{configuration}}(\langle \mathit{SBJ}, \mathit{PRD}, \mathit{OBJ}, \mathit{COM}\rangle \mid \{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\}@\mathrm{S})}$

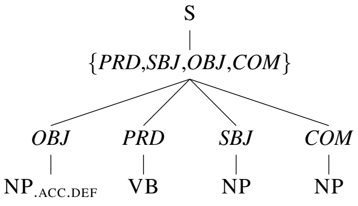
$\mathbf{P}_{\text{realization}}(\mathrm{VB} \mid \mathit{PRD}@\mathrm{S})$

$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{SBJ}@\mathrm{S})$

$\mathbf{P}_{\text{realization}}(\mathrm{NP}_{.\text{ACC.DEF}} \mid \mathit{OBJ}@\mathrm{S})$

$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{COM}@\mathrm{S})$

(5b)



⇒ $\mathbf{P}_{\text{projection}}(\{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\} \mid \mathrm{S})$

$\boxed{\mathbf{P}_{\text{configuration}}(\langle \mathit{OBJ}, \mathit{PRD}, \mathit{SBJ}, \mathit{COM}\rangle \mid \{\mathit{PRD}, \mathit{SBJ}, \mathit{OBJ}, \mathit{COM}\}@\mathrm{S})}$

$\mathbf{P}_{\text{realization}}(\mathrm{VB} \mid \mathit{PRD}@\mathrm{S})$

$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{SBJ}@\mathrm{S})$

$\mathbf{P}_{\text{realization}}(\mathrm{NP}_{.\text{ACC.DEF}} \mid \mathit{OBJ}@\mathrm{S})$

$\mathbf{P}_{\text{realization}}(\mathrm{NP} \mid \mathit{COM}@\mathrm{S})$

Figure 7: Differential Object-Marking (DOM)

	Model	*Plain*	*Head*	*Parent*	*ParentHead*
Base					
	SP-PCFG	67.61/68.77	71.01/72.48	73.56/73.79	73.44/73.61
	RR-PCFG	65.86/66.86	71.84/72.76	74.06/74.28	75.13/75.29
BaseDef					
	SP-PCFG	67.68/68.86	71.17/72.47	74.13/74.39	72.54/72.79
	RR-PCFG	66.65/67.86	73.09/74.13	74.59/74.59	76.05/76.34
BaseDefAcc					
	SP-PCFG	68.11/69.30	71.50/72.75	74.16/74.41	72.77/73.01
	RR-PCFG	67.13/68.01	73.63/74.69	74.65/74.79	**76.15/76.43**

Table 1: Differential Object Marking: F1 for sentences 1–500 in the Treebank. *Plain, Parent, Head* are syntactic splits. *Base, Def, Acc* are morphological splits.

4 Evaluation

Tsarfaty and Sima'an (2008), Tsarfaty et al. (2009) and Tsarfaty (2010) report a series of experiments that learn RR descriptions from treebank data and use them for wide coverage statistical parsing of Modern Hebrew. Here we limit the discussion to the methodological outline and to highlighting the main results. Interested readers are encouraged to follow up on the detailed analysis in the original articles.

All of the reported experiments use data from the Modern Hebrew treebank. The models are trained on sentences 500-5500 and tested on sentences 1-500. An automatic procedure is used to read off RR parameters from phrase-structure trees augmented with functional and morphological features. The paradigmatic representation of constituents uses the treebank labels' set, morphological information, and the relation labels' set {SBJ,PRD,OBJ, COM, I-COM,CNJ}. The lexical category of the predicate is percolated to each syntactic constituent in the representation. The training procedure uses simple relative frequency estimates and rare-words distribution for lexical smoothing. A general purpose CKY implementation is used for parsing, and all experiments are evaluated using Parseval on trees in canonical form (i.e., all RR-specific information is removed prior to evaluation).

Tsarfaty and Sima'an (2008) show that RR versions of the treebank grammar perform at the same level or significantly better than PCFGs that use history-based conditioning context. Moreover, morphological information is seen to contribute significant improvements for an RR treebank grammar, while leading to performance degradation with other PCFGs, as recapitulated in table 1. Tsarfaty et al. (2009) show that an RR grammar augmented with differential-object marking features significantly outperforms different versions of Head-Driven treebank grammars à la Collins (2003). The RR grammars in Tsarfaty et al. (2009) are more economic than Head-Driven ones learned for the same set of data. Both Tsarfaty and Sima'an (2008) and Tsarfaty et al. (2009) guess the PoS tags of words. Tsar-

Model		*Base*	*BaseGen*	*BaseDefAcc*	*BaseGenDefAcc*
SP-AGR	*Plain*	79.77 (3942)	79.55 (7594)	80.13 (4980)	*80.26* *(8933)*
RR-AGR	*Plain*	80.23 (3292)	81.09 (5686)	81.48 (3772)	**82.64** **(6516)**
SP-AGR	*Parent*	*83.06* *(5914)*	82.18 (10765)	79.53 (12700)	80.89 (11028)
RR-AGR	*Parent*	83.49 (6688)	83.70 (10063)	83.66 (12383)	**84.13** **(12497)**
SP-AGR	*Parent* *Head*	*76.61* *(10081)*	64.07 (16721)	75.12 (11681)	61.69 (18428)
RR-AGR	*Parent* *Head*	**83.40** **(12497)**	81.19 (22979)	83.33 (13828)	80.45 (24934)

Table 2: Differential Object Marking and Agreement for gold PoS tagged input. *Plain, Parent, Head* are syntactic splits. *Base, Def, Acc* are morphological splits.

faty (2010) reports parsing results for an extended set of features, including DOM features and gender agreement, when parsing gold PoS-tagged input. These results are summarized in table 2. The best result here (F_1 84.13) constitutes the best parsing result reported so far for Hebrew in the gold PoS-tags setting.

Tsarfaty (2010) finally shows that the parameter tables read off from the treebank can provide an immediate probabilistic interpretation for typological descriptions of the language. For instance, a probability distribution over production probabilities at the left of table 3 confirms the observation that Hebrew is primarily an SVO language, while allowing for word-order variation. The probability distribution over the realization of objects captures, for different types of lexical heads, the DOM pattern discussed in section 3, with a sharp distribution. Probability tables showing sharp distributions for morphological realization parameters and less sharp distributions for configuration parameters, reflect the *less-configurational* nature of Hebrew. If we are to estimate the probability distributions of RR parameters for different languages, comparing the empirical distributions we obtain may provide us with a precise way to quantify different levels of *nonconfigurationality*.

5 Conclusion

The idea presented here, viewing syntactic categories as designating paradigms and augmenting them with realization rules, provides for a powerful modeling strategy which can be developed into a complete architecture of specifying and statistically learning syntactic descriptions. The Relational-Realizational (RR) architecture developed herein is particularly adequate for languages that exhibit rich morphosyntactic interactions. The RR architecture is simple in the sense that it alternates three

Probability	*Configuration*	
1 %	□ *PRD* □ *SBJ OBJ* □	VSO
1.3%	*SBJ* □ *PRD OBJ* □	SVO
1.7%	□ *PRD OBJ SBJ* □	VOS
1.7%	□ *SBJ PRD* □ *OBJ* □	SVO
3%	*OBJ PRD SBJ* □	OVS
3.7%	□ *PRD SBJ* □ *OBJ* □	VSO
4.1%	*SBJ* □ *PRD* □ *OBJ* □	SVO
6.5%	□ *SBJ PRD OBJ* □	SVO
10.3%	*SBJ* □ *PRD OBJ* □	SVO
12.3%	□ *PRD SBJ OBJ* □	VSO
15.6%	*SBJ PRD* □ *OBJ* □	SVO
35.3%	*SBJ PRD OBJ* □	SVO

Probability	*Realization*
5.8%	$\text{NP}_{\text{DEF.ACC}}\langle\text{PRP}\rangle$@S
6.5%	$\text{NP}_{\text{DEF.ACC}}\langle\text{NNT}\rangle$@S
6.7%	$\text{NP}_{\text{DEF.ACC}}\langle\text{NN}_{\text{DEF}}\rangle$@S
7.4%	$\text{NP}_{\text{DEF.ACC}}\langle\text{NNP}\rangle$@S
8.8%	$\text{NP}\langle\text{NNT}\rangle$@S
14.7%	$\text{NP}_{\text{DEF.ACC}}\langle\text{NN}\rangle$@S
43.5%	$\text{NP}\langle\text{NN}\rangle$@S

Table 3: Word-Order and Object-Marking Parameter Tables (**P**(x)$> 1\%$)

phases of generation for each constituent, it can be specified in a fully formal way, it is robust in the sense that it can be easily applied to treebank data, and it can be used to automatically learn treebank grammars for efficient and accurate parsing. The probabilistic parameters learned by the model are easily interpretable as indicating morphological and structural dimensions of typological variation, which can potentially be developed into empirical measures of the level of nonconfigurationality of different languages. At the same time, the models presented here use a particular set of independence assumptions which conceptualizes morphology as orthogonal to positions. Future versions will also explore different assumptions, for learning rules that spell-out the morphological and syntactic realization jointly.

References

Aissen, Judith. 2003. Differential Object Marking: Iconicity vs. Economy. *Natural Language and Linguistic Theory* 49, 435–483.

Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge.

Aronoff, Mark. 1994. *Morphology By Itself*. Cambridge: The MIT Press.

Blevins, James P. 2010. *Word and Paradigm Morphology*. Oxford University Press.

Bloomfield, Leonard. 1933. *Language*. Holt, Rinehart and Winston Inc.

Bresnan, Joan. 2000. *Lexical-Functional Syntax*. Blackwell.

Cahill, Aoife, Burke, Michael, O'Donovan, Ruth, Riezler, Stefan, van Genabith, Josef and Way, Andy. 2008. Wide-Coverage Deep Statistical Parsing using Automatic Dependency Structure Annotation. *Computational Linguistics* 34(1).

Charniak, Eugene. 1996. Tree-Bank Grammars. In *AAAI/IAAI, Vol. 2*.

Collins, Michael. 2003. Head-Driven Statistical Models for Natural Language Parsing. *Computational Linguistics* 29(4).

Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Joseph H. Greenberg (ed.), *Universals of Language*, pages 73–113, MIT Press.

Hockenmaier, Julia and Steedman, Mark. 2002. Acquiring Compact Lexicalized Grammars from a Cleaner Treebank. In *Proceedings of LREC*.

Kübler, Sandra. 2008. The PaGe Shared Task on Parsing German. In *Proceedings of the ACL Workshop on Parsing German*.

Matthews, Peter H. 1974. *Morphology*. Cambridge University Press.

Matthews, Peter H. 1981. *Syntax*. Cambridge University Press.

Melnik, Nurit. 2002. Verb-Initial Constructions in Modern Hebrew. Ph.D. Thesis. University of Califiornia at Berkeley.

Miyao, Yusuke and Tsujii, Jun'ichi. 2008. Feature-Forest Models for Probabilistic HPSG Parsing. *Computational Linguistics* 34(1), 35–80.

Pike, Kenneth L. 1963. A Syntactic Paradigm. *Language* 39(2), 216–230.

Postal, Paul M. and Perlmutter, David M. 1977. Toward a Universal Characterization of Passivization. In *BLS 3*.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. Brace and company.

Shlonsky, Ur and Doron, Edit. 1991. Verb Second in Hebrew. In Dawn Bates (ed.), *The Proceedings of the Tenth West Coast Conference on Formal Linguistics*.

Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge Studies in Linguistics, No. 93, Cambridge University Press.

Tsarfaty, Reut. 2010. Relational-Realizational Parsing. Ph.D. Thesis. University of Amsterdam.

Tsarfaty, Reut, Seddah, Djame, Goldberg, Yoav, Kübler, Sandra, Candito, Marie, Foster, Jenifer, Versley, Yannick, Rehbein, Ines and Tounsi, Lamia. 2010. Statistical Parsing of Morphologically Rich Languages: What, How and Whither. In *Proceedings of NA-ACL workshop on Parsing Morphologically Rich Languages*.

Tsarfaty, Reut and Sima'an, Khalil. 2008. Relational-Realizational Parsing. In *Proceedings of CoLing*.

Tsarfaty, Reut, Sima'an, Khalil and Scha, Remko. 2009. An Alternative to Head-Driven Approaches for Parsing a (Relatively) Free Word Order Language. In *Proceedings of EMNLP*.