Modeling Morphosyntactic Agreement in Constituency-Based Parsing of Modern Hebrew

Reut Tsarfaty* **and Khalil Sima'an**
Institute for Logic, Language and Computation
University of Amsterdam
{r.tsarfaty,k.simaan}@uva.nl

Abstract

We show that naïve modeling of morphosyntactic agreement in a Constituency-Based (CB) statistical parsing model is worse than none, whereas a linguistically adequate way of modeling inflectional morphology in CB parsing leads to improved performance. In particular, we show that an extension of the Relational-Realizational (RR) model that incorporates agreement features is superior to CB models that treat morphosyntax as state-splits (SP), and that the RR model benefits more from inflectional features. We focus on parsing Hebrew and report the best result to date, $F_1$84.13 for parsing off of gold-tagged text, 5% error reduction from previous results.

1 Introduction

Agreement is defined by linguists as the systematic covariance of the grammatical properties of one linguistic element to reflect the semantic or formal properties of another (Corbett, 2001). Morphologically marked agreement features such as *gender, number* and *person* are used to realize grammatical relations between syntactic constituents, and such patterns are abundantly found in (less- or) non-configurational languages (Hale, 1983) where the order of words is known to be (relatively) free. Agreement features encompass information concerning the functional relations between constituents in the syntactic structure, but whether incorporating agreement features in a statistical parsing model leads to improved performance has so far remained an open question and saw contradictory results.

*The first author is currently a researcher at the department of Linguistics and Philology at Uppsala University.

Taking Semitic languages as an example, it was shown that an SVM-based shallow parser (Goldberg et al., 2006) does not benefit from including agreement features for NP chunking in Hebrew. Phrase-structure based parsers for Arabic systematically discard morphological features from their label-set and never parametrize agreement explicitly (Maamouri et al., 2008). Models based on deep grammars such as CCG (Hockenmaier and Steedman, 2003) and HPSG (Miyao and Tsujii, 2008) could in principle use inflectional morphology, but they currently rely on functional information mainly. For formalisms that do incorporate morphology, generative models are may leak probability due to unification failures (Abney, 1997). Even results from dependency parsing remain inconclusive. It was shown for dependency parsing that case, definiteness and animacy features are useful to enhance parsing (e.g., (Øvrelid and Nivre, 2007)), agreement patterns are often excluded. When agreement features were included as features in dependency parser for Hebrew in (Goldberg and Elhadad, 2009) for Hebrew they obtained tiny-to-no improvement.

A question thus emerges whether there are any benefits in explicitly incorporating morphosyntactic agreement patterns into our models. This question is a manifestation of a greater issue, namely, whether it is beneficial to represent complex patterns of morphology in the statistical parsing model, or whether configurational information subsume the relevant patterns, as it is commonly assumed in constituency-based parsing. Here we claim that agreement features are useful for statistical parsing provided that they are represented and parametrized in a way that reflects their linguistic substance; to express functional information orthogonal to configuration.

Proceedings of the NAACL HLT 2010 First Workshop on Statistical Parsing of Morphologically-Rich Languages, pages 40–48,
Los Angeles, California, June 2010.

We do so by extending the Relational-Realizational (RR) model we presented in (Tsarfaty and Sima'an, 2008) to explicitly encode agreement features in its native representation (RR-AGR). In the RR model, a joint distribution over grammatical relations is firstly articulated in the *projection* phase. The grammatical relations may be spelled out by positioning them with respect to one another in the *configuration* phase, through the use of morphology in the *realization* phase, or both. This paper shows that, for Hebrew, this RR-AGR strategy significantly outperforms a constituency-based model that treats agreement features as internally structured non-terminal state-splits (SP-AGR). As we accumulate morphological features, the performance gap between the RR and SP models becomes larger.

The best result we report for the RR-AGR model, $F_1$84.13, is the best result reported for Hebrew to date for parsing gold PoS-tagged segments, with 5% error reduction from previous results. This result is also significantly higher than all parsing results reported so far for Arabic, a Semitic language with similar morphosyntactic phenomena.[1] The RR approach is shown to be an adequate way to model complex morphosyntactic patterns for improving constituency-based parsing of a morphologically rich, free word order language. Because the RR model is also proper and generative, it may also embed as a language model to enhance more complex NLP tasks, e.g., statistical Machine Translation.

2 The Data

The grammar of nonconfigurational languages allows for freedom in word ordering and discontinuities of syntactic constituents (Hale, 1983). Such languages do not rely on configurational information such as position and adjacency in marking grammatical relations such as *subject* and *object*, but instead they use word-level morphology. One way to encode grammatical relations in the form of words is by using morphological case, that is, explicitly marking an argument (e.g. nominative, accusative) with respect to its grammatical function. In (Tsarfaty et al., 2009) we showed that incorporating case indeed leads to improved performance for constituency-based, Relational-Realizational parsing of Hebrew.

[1] In (Maamouri et al., 2008), $F_1$78.1 for gold standard input.

A more involved way to morphologically encode grammatical relations is by making explicit reference to the properties of multiple linguistic elements. This is the general pattern of agreement, i.e.,

> "[A] systematic covariance between a semantic or a formal property of one element and a formal property of another." (Steele, adapted from (Corbett, 2001))

Describing agreement patterns involves explicit reference to the following four components; the element which determines the agreement properties is the **Controller** of the agreement, the element whose properties are determined by agreement is the **Target**, the syntactic environment in which the agreement occurs is the **Domain** of agreement, and the properties with respect to which they agree are agreement **Features** (Corbett, 2001). Agreement is an inherently asymmetrical relation. Combination of features displayed by controllers has to be accommodated by the inflectional features of the target, but there is no opposite requirement. Let us illustrate the formal description of agreement through Subject-Verb agreement familiar from English (1).

(1) a. Subject-Verb Agreement in English:
Controller: NP
Target: V
Domain: S
Features: *number, person*
b. Example:
i. They like the paper
ii. *They likes the paper

The agreement target (the verb) in English has a rich enough inflectional paradigm that reflects the person and number features inherent in controllers — the nouns that realize subjects. (But nouns in English need not reflect, say, tense.) Had the subject been an NP, e.g., the phrase "the committee", the agreement pattern would have had to be determined by the features of the entire NP, and in English the features of the phrase would be determined by the lexical head "committee". The controller of the agreement (noun) does not coincide with the head of the lexical dependency (the verb), which means that the direction of morphological dependencies need not coincide with that of lexical dependencies.

The Semitic Language Modern Hebrew Modern Hebrew, (henceforth, Hebrew) is a Semitic language with a flexible word order and rich morphological structure. Hebrew nouns morphologically reflect their inherent *gender* and *number*. Pronouns also reflect *person* features. Hebrew verbs are inflected to reflect *gender, number, person* and *tense*. Adjectives are inflected to reflect the inherent properties of nouns, and both nouns and adjectives are inflected for *definiteness*. The Hebrew grammar uses this arsenal of properties to implement a wide variety of agreement patterns realizing grammatical relations.

Agreement in Hebrew S Domains Hebrew manifests different patterns of agreement in its S domain. Verbal predicates (the target) in matrix sentences (the domain) agree with their nominal subjects (the controller) on the agreement features *gender, number* and *person*. This occurs regardless of their configurational positions, as illustrated in (2b).

(2) a. Agreement in Verbal Sentences:
Controller: NP
Target: V
Domain: S
Features: *number, person, gender*

b. i. דני נתן מתנה לדינה

dani natan matana
Dani.3MS gave.3MS present
ledina
to-Dina

Dani gave a present to Dina (SVO)

ii. מתנה נתן דני לדינה

matana natan dani
present gave.3MS Dani.3MS
ledina
to-Dina

Dani gave a present to Dina (VI)

Subject-Predicate agreement relations are not only independent of surface positions, but are also orthogonal to the syntactic distributional type of the constituent realizing the predicate. Semitic languages allow for predicates to be realized as an NP, an ADJP or a PP clause (3b) lacking a verb altogether. (In the Hebrew treebank, such predicates are marked as PREDP). In all such cases, agreement feature-bundles realized as pronominals, which (Doron, 1986) calls Pron, are optionally placed after the subject. The position of Pron element with respect to the subject and predicate is fixed.[2] The role of these Pron elements is to indicate the argument-structure of a nominal sentence that is not projected by a verb. In the Hebrew treebank they are subsumed under predicative phrases (PREDPs). If a PREDP head is of type NP or ADJP it must be inflected to reflect the features of the subject controller, as is illustrated in examples (3b-i)–(3b-ii).

(3) a. Agreement in Nominal Sentences:
Controller: NP
Target: Pron
Domain: S
Features: *number, gender,person*

b. i. דינה (היא) ציירת

dina (hi) cayeret
Dina.FS (Pron.3FS) painter.FS

Dina is a painter

ii. דינה (היא) מוכשרת

Dina (hi) muchsheret
Dina.FS (Pron.3FS) talented.FS

Dina is talented

iii. דינה (היא) בבית

Dina (hi) babayit
Dina.FS (Pron.3FS) in-the-house

Dina is at home

c. i. (היא)* דינה ציירת (hi)* dina cayeret
(Pron.3FS)* Dina.FS painter.FS

The pronominal features *gender, number, person* are also a part the inflectional paradigm of the verb היה (be), which is extended to include tense features. These inflected elements are used as AUX which function as co-heads together with the main (nominal or verbal) predicate. AUX elements that take a nominal predicate as in (4b) agree with their subject, and so do auxiliaries that take a verbal complement, e.g., the modal verb in (4c). The nominal predicate in (4b) also agrees with the subject – and so does the modal verb in (4c). Agreement of AUX with the

[2]Doron (1986) shows that these Pron elements can not be considered the present tense supplements of AUX elements in Hebrew since their position with respect to the subject and predicate is fixed, whereas AUX can change position, see (4) below.

verbal or nominal predicates is again independent of their surface positions.

(4) a. Subject-AUX Agreement in Hebrew:
Controller: NP
Target: AUX
Domain: S
Features: *number, person, gender*

b. i. היא היתה בעבר ציירת
hi hayta be'avar
she.3FS was.3FS in-past
cayeret
painter.FS
She was a painter in the past

ii. בעבר היתה היא ציירת
be'avar hayta hi
in-past was.3FS she.3FS
cayeret
painter.FS
She was a painter in the past"

c. i. היא היתה אמורה להגיע
hi hayta amura
She.3FS was.3FS supposed.FS
lehagi'a
to-arrive
She was supposed to arrive

ii. היא אמורה היתה להגיע
hi amura hayta
She.3FS supposed.FS was.3FS
lehagi'a
to-arrive
She was supposed to arrive

Agreement in Construct State Nouns Semitic languages allow for the creation of noun compounds by phonologically marking their lexical head and adding a genitive complement. These constructions are called Construct-State Nouns (CSN) (Danon, 2008) and an example of a CSN is provided in (5a).[3]

(5) a. בת הצייר
bat ha-cayar
child.FS.CSN Def-painter.MS
The painter's daughter

[3]Also known as *iDaFa* constructions in Arabic.

In such cases, all the agreement features are taken from the head of the CSN, the noun 'daughter' in (5). Since CSNs may be embedded in other CSNs, the constructions may be arbitrarily long. When short or long, CSNs themselves may be modified by adjectives that agree with the CSN as a whole. This gives rise to multiple patterns of agreement within a single complex CSN. Consider, for instance, the modified CSN in (6a).

(6) a. בת הצייר המוכשרת
bat ha-cayar
child.FS.CSN Def-painter.MS
ha-muchsheret
Def-talented.FS
The talented daughter of the painter

The features Def, F, S of the adjective 'talented' agree with the inherent properties of the CSN head 'child.FS' and with the definiteness status of the embedded genitive Def-painter. This phenomenon is called by Danon (2008) definiteness-spreading, and what is important about such spreading is to observe that it is not always the case that all agreement features of a phrase are contributed by its lexical head.[4]

Interim Summary The challenges of modeling agreement inside constituency-based statistical models can be summarized as follows. The models are required to assign probability mass to alternating sequences of constituents while retaining equivalent feature distributions that capture agreement. Agreement is (i) orthogonal to the position of constituents (ii), orthogonal to their distributional types, and (iii) orthogonal to features' distributions among dominated subconstituents. Yet, from a functional point of view their contribution is entirely systematic.

3 The Models

The strong version of the well-known Lexicalist Hypothesis (LH) states that "syntactic rules cannot make reference to any aspect of word internal structure" (Chomsky, 1970). Anderson (1982) argues that syntactic processes operating within configurational structures can often manipulate, or have access to, formal and inherent properties of individual words. Anderson (1982) argues that a model

[4]Examples for non-overlapping contribution of features by multiple dependencies can be found in (Guthmann et al., 2009).

that is well-equipped to capture such phenomena is one that retains a relaxed version of the LH, that is, one in which syntactic processes do not make reference to aspects of word-internal structure *other than morphologically marked inflectional features*. What kind of parsing model would allow us to implement this relaxed version of the Lexicalist Hypothesis?

The Morphosyntatctic State-Splits (SP) Model One way to maintain a relaxed version of the LH in syntax is to assume a constituency-based representation in which the morphological features of words are percolated to the level of constituency in which they are syntactically relevant. This approach is characteristic of feature-based grammars (e.g., GPSG (Gazdar et al., 1985) and follow-up studies). These grammars assume a feature geometry that defines the internal structure of node labels in phrase-structure trees.[5]

Category-label state-splits can reflect the different morphosyntactic behavior of different non-terminals of the same type. Using such supervised, linguistically motivated, state-splits, based on the phrase-level marking of morphological information is one may build an efficient implementation of a PCFG-based parsing model that takes into account morphological features. State-split models were shown to obtain state-of-the-art performance with little computational effort. Supervised state-splits for constituency-based unlexicalized parsing in (Klein and Manning, 2003) in an accurate English parser. For the pair of Hebrew sentences (2b), the morphological state-split context-free representation of the domain S is as described at the top of figure 1.[6]

The Relational-Realizational (RR) Model A different way to implement a syntactic model that conform to the relaxed LH is by separating the inflectional features of surface words from their grammatical functions in the syntactic representation and letting the model learn systematic form-function correspondence patterns between them.

The Relational-Realizational (RR) model (Tsarfaty and Sima'an, 2008) takes such a 'separationalist' approach which is constituent-based. Grammatical relations are separated from their morphological or syntactic means of realization, which are in turn also distinguished. The easiest way to describe the RR model is via a three-phase generative process encompassing the *projection, configuration* and *realization* phases. In the *projection* phase, a clause-level syntactic category generates a Relational Network (RN), i.e., a set of grammatical function-labels representing the argument-structure of the clause. In the *configuration* phase, linear ordering is generated for the function-labels and optional realization slots are reserved for elements such as punctuation, auxiliaries and adjuncts. The *realization* phase spells out a rich morphosyntactic representation (MSR) — a syntactic label plus morphological features — realizing each grammatical function and each of the reserved slots. The process repeats as necessary until MSRs of pre-terminals are mapped to lexical items.

In (Tsarfaty et al., 2009) we have shown that the RR model makes beneficial use of morphological patterns involving case marking, but did not study the incorporation of inflectional agreement features such as *gender*. Since agreement features such as *gender, number* and case-related information such *accusativity, definiteness* are determined by non-overlapping subconstituents, it remains an open question whether an addition of agreement features into the model can be down in a linguistically adequate and statistically sound way, and whether or not they further improve performance.

We claim that the Relational-Realizational model of (Tsarfaty et al., 2009) has all the necessary ingredients to seamlessly migrate RR representations to ones that encode agreement explicitely. In order to explain how we do so let us recapitulate the empirical facts. Agreement is an asymmetric relation defined for a certain domain, in which the agreement properties of a target co-vary with the inherent properties of the controller. Consider the two sentences in (2b) in which the formal means to differentiate the subject from the object is by the pattern of an agreeing predicate. The RR representations of the domain S are given at the bottom of figure 1.

[5]While agreement patterns in feature-rich grammars give rise to re-entrancies that break context-freeness, GPSG shows that using feature-percolation we can get quite far in modeling morphosyntactic dependencies and retaining context-freeness.

[6]Horizontal markovization à la (Klein and Manning, 2003) would be self-defeating here. Markovization of constituents conditions inflectional features on configurational positions, which is inadequate for free word-order languages as Hebrew. This is already conjectured in the PhD thesis of Collins, and it is verified empirically for Hebrew in (Tsarfaty et al., 2009).

The agreement targets and agreement controllers are easy to recognize; controllers are the syntactic constituents that realize subjects, parametrized as $\mathbf{P}_{\text{realization}}(VB|PRD@S)$, and targets are the ones that realize predicates, parametrized as $\mathbf{P}_{\text{realization}}(NP|SBJ@S)$. Now, if we take the predicted labels of controllers and targets to include reference to inflectional features, we get the following parameterization of the realization parameters $\mathbf{P}_{\text{realization}}(VB\langle\text{FEATS}_i\rangle|PRD@S)$ and $\mathbf{P}_{\text{realization}}(NP\langle\text{FEATS}_j\rangle|SBJ@S)$ with $\langle\text{FEATS}_i\rangle$, $\langle\text{FEATS}_j\rangle$ the inflectional features indicated in their morphosyntactic representation. Now, we only need to make sure that $\langle\text{FEATS}_i\rangle$, $\langle\text{FEATS}_j\rangle$ indeed agree, *regardless of their position under S*.

We do so by explicitly marking the domain of agreement, the S category, with the features of the syntactically most prominent participant in the situation, the subject (this is where the non-symmetrical nature of agreement comes into play). The realization distributions take the following forms $\mathbf{P}_{\text{realization}}(VB\langle\text{FEATS}_j\rangle|PRD@S\langle\text{FEATS}_i\rangle)$ and $\mathbf{P}_{\text{realization}}(NP\langle\text{FEATS}_i\rangle|SBJ@S\langle\text{FEATS}_i\rangle)$. In the former, $NP\langle\text{FEATS}_i\rangle$ reflects the inherent features of the SBJ and in the latter $VB\langle\text{FEATS}_j\rangle$ reflects the agreement features of the PRD. Now, regardless of word order, and regardless of the internal structure of NPs, the parameters capturing agreement would be the same for examples (2b i-ii). The only parameters that differ are the configuration parameters (boxed), reflecting word-order alternation.

For the sake of completeness we include here also the SP vs. RR representation of S domains involving auxiliaries in figure 2. Here the sentences vary only in the position of the AUX element relative to the subject with which it agrees. Subjects, predicates, and slots that have been reserved for AUX elements, all reflect the same pattern of agreement through their conditioning on the rich representation of the domain.[7] More parameters that vary here (boxed) are AUX placement and realization parameters. Since Pron elements endow PREDPs with agreement features, agreement with verbless (nominal) predicates under S analogously follows.

[7]In Hebrew, even some adverbial modifiers reflect patterns of agreement, e.g., עודני (literally, "I am still", glossed 'still.1S'). This solution caters for all such patterns in which non-obligatory elements exhibit agreement.

4 Experiments

We aim to examine whether the explicit incorporation of agreement features helps Hebrew parsing, and if so, which of the two modeling strategies is better for utilizing the disambiguation cues provided by morphosyntactic agreement.

Data We use the Hebrew treebank v2.0 with the extended annotation of (Guthmann et al., 2009), which adds inflectional properties to non-terminal categories such as NP and VP. We head-annotate the corpus and systematically add the agreement features of Domains throughout the treebank. We further distinguish finite from non-finite verb forms, and cliticized from non-cliticized nouns, as in (Goldberg and Tsarfaty, 2008; Tsarfaty et al., 2009). On top of the treebank labels SBJ subject, OBJ object, COM complement and CNJ conjunction we add PRD predicates and IC infinitival complements.

Procedure We devised a procedure to read-off treebank grammars based on (i) GPSG-like, statesplit context-free parameters (SP-AGR), and (ii) RR-AGR parameters in which context-free rules capture the *projection, configuration* and *realization* phases. In each model the multiplication provides the probability of the generation. We use relative frequency estimates and exhaustively parse gold pos-tagged input[8] using a general-purpose CKY parser. We use the same data split as in (Goldberg and Tsarfaty, 2008; Tsarfaty et al., 2009) (training on sentences 501-6000 and parsing sentences 1-500) and we convert all trees to the flat, coarse-grained, original treebank representation for the purpose of evaluation.

Setup We experiment with bare constituent labels, grand-parent decorated labels (gp), and labels decorated with grand-parent and head-tag labels (gp,hd). We use increasingly richer subsets of the $\{$*gender, definiteness, accusativity*$\}$ set.[9]

[8]This choice to parse gold-tagged sentences is meant to alleviate the differences in the model's morphological disambiguation capacity. We want to evaluate the contribution of morphological features for syntactic disambiguation, and if the models will disambiguate the morphological analyses differently, the syntactic analysis will be assigned to different yields and the accuracy results would be strictly incomparable. But see (Goldberg and Tsarfaty, 2008) for a way to combine the two.

[9]We deliberately choose features that have non-overlapping behavior, to see whether their contribution is accumulative.

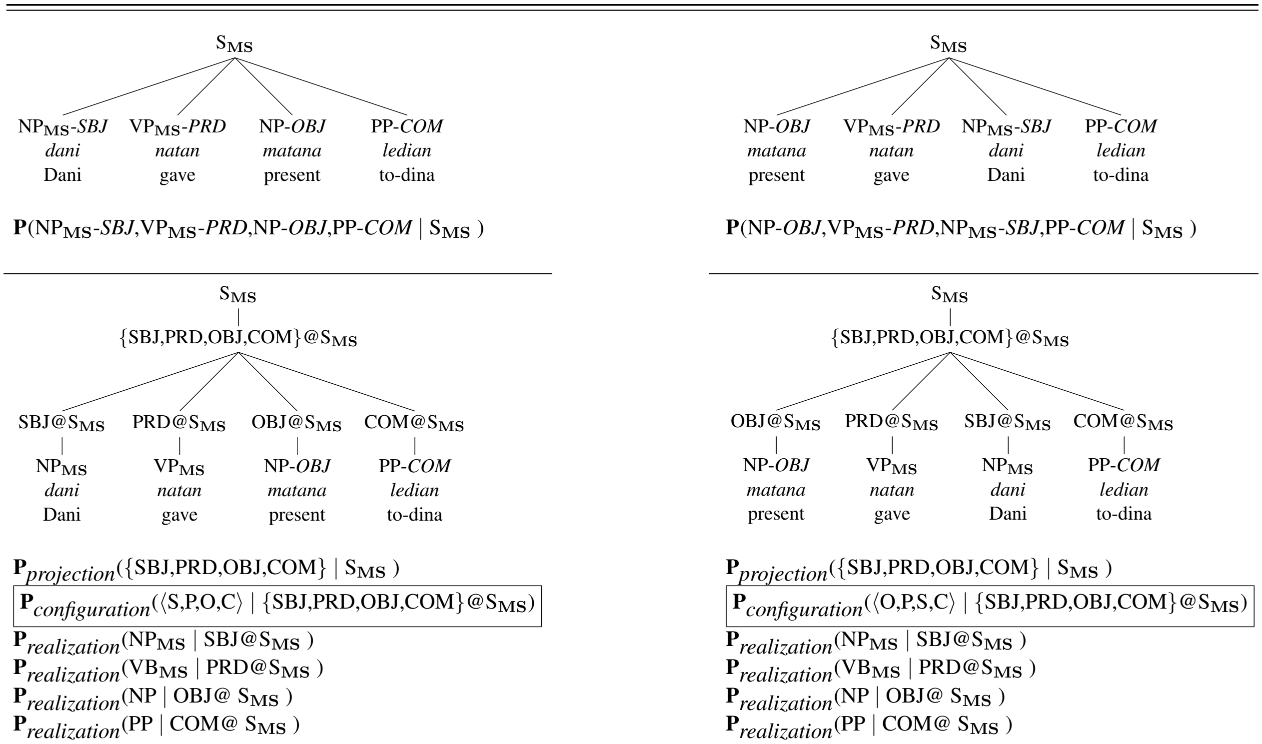

Figure 1: The SP-AGR (top) and RR-AGR representations of sentences (2b-i) (left) and (2b-ii).

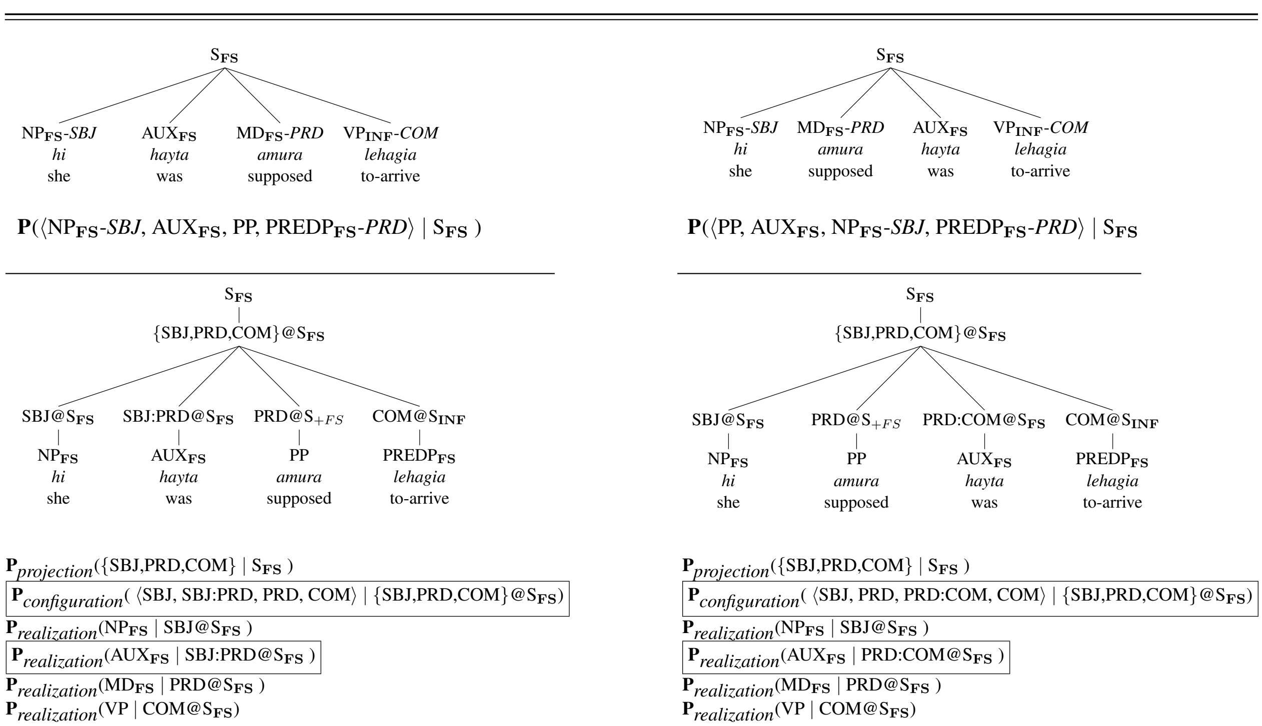

Figure 2: The SP-AGR (top) and RR-AGR representation of sentences (4c-i) (left) and (4c-ii).

Model	∅	gender	def+acc	gender+def+acc
SP-AGR	79.77	79.55	80.13	80.26
	(3942)	(7594)	(4980)	(8933)
RR-AGR	80.23	81.09	81.48	**82.64**
	(3292)	(5686)	(3772)	**(6516)**
SP-AGR (gp)	83.06	82.18	79.53	80.89
	(5914)	(10765)	(12700)	(11028)
RR-AGR (gp)	83.49	83.70	83.66	**84.13**
	(6688)	(10063)	(12383)	**(12497)**
SP-AGR (gp,hd)	76.61	64.07	75.12	61.69
	(10081)	(16721)	(11681)	(18428)
RR-AGR (gp,hd)	**83.40**	81.19	83.33	80.45
	(12497)	(22979)	(13828)	(24934)

Table 1: F-score (#params) measure for all models on the Hebrew treebank dev-set for Sentences Length < 40

5 Results and Discussion

Table 1 shows the standard F1 scores (and #parameters) for all models. Throughout, the RR-AGR model outperforms the SP-AGR models that use the same category set and the same morphological features as state splits. For RR-AGR and RR-AGR (gp) models, adding agreement features to case features improves performance. The accumulative contribution is significant. For SP-AGR and SP-AGR (gp) models, adding more features either remains at the same level of performance or becomes detrimental.

Since the SP/RR-AGR and SP/RR-AGR (gp) models are of comparable size for each feature-set, it is unlikely that the differences in performance are due to the lack of training data. A more reasonable explanation if that the RR parameters represent functional generalizations orthogonal to configuration for which statistical evidence is more easily found in the data. The robust realization distributions which cut across ordering alternatives can steer the disambiguation in the right direction.

The RR-AGR (gp) +gen+def+acc model yields the best result for parsing Hebrew to date (F1 84.13), improving upon our best model in (Tsarfaty et al., 2009) (F1 83.33, underlined) in a pos-tagged setting. For this setting, Arabic parsing results are F1 78.1. Given the similar morphosyntactic phenomena (agreement, MaSDaR, iDaFa) it would be interesting to see if the model enhances parsing for Arabic.

For (gp,hd) models (a configuration which was shown to give the best results in (Tsarfaty et al., 2009)) there is a significant decrease in accuracy with the gender feature, but there is a lesson to be learned. Firstly, while the RR-AGR (gp,hd) model shows moderate decrease with gender, the decrease in performance of SP-AGR (gp,hd) for the same feature-set is rather dramatic, which is consistent with the observation that the RR model is less vulnerable to sparseness and that it makes better use of the statistics of functional relations in the data.

Consulting the size of the different grammars, the combination of RR-AGR (gp, hd) with gender features indeed results in substantially larger grammars, and it is possible that at this point we indeed need to incorporate smoothing. At the same time there may be an alternative explanation for the decreased performance. It might be that the head-tag does not add informative cues beyond the contribution of the features which are spread inside the constituent, and are already specified. This is a reasonable hypothesis since *gender* in Hebrew always percolates through the head as opposed to *def/acc* that percolate from other forms. Incorporating head-tag in (Tsarfaty et al., 2009) might have led to improvement only due to the lack of agreement features which subsume the relevant pattern. This suggests that incorporating all co-heads and functional elements that contribute morphological features spread inside the constituent, is more adequate for modeling morphosyntax than focusing on the features of a single head.

6 Conclusion

We show that morphologically marked agreement features can significantly improve parsing performance if they are represented and parametrized in a way that reflects their linguistic substance: relating form-and-function in a non-linear fashion. We have so far dealt with the adequacy of representation and we plan to test whether more sophisticated estimation (e.g., split-merge-smooth estimation as in (Petrov et al., 2006)) can obtain further improvements from the explicit representation of agreement. At the same time, the state-of-the-art results we present render the RR model promising for further exploration with morphologically rich languages.

Acknowledgements The work of the first author has been funded by NWO, grant 017.001.271. We wish to thank Joakim Nivre and three anonymous reviewers for helpful comments on earlier drafts.

References

Steven Abney. 1997. Stochastic attribute-value grammars. *Computational Linguistics*, 23(4):597–618.

Stephen R. Anderson. 1982. Where's morphology? *Linguistic Inquiry*.

Noam Chomsky. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum, editors, *Reading in English Transformational Grammar*. Waltham: Ginn.

Greville G. Corbett. 2001. Agreement: Terms and boundaries. In *SMG conference papers*.

Gabi Danon. 2008. Definiteness spreading in the hebrew construct-state. *Lingua*, 118(7):872–906.

Edit Doron. 1986. The pronominal "copula" as agreement clitic. *Syntax and Semantics*, (19):313–332.

Gerald Gazdar, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag. 1985. *Generalised phrase structure grammar*. Blackwell, Oxford, England.

Yoav Goldberg and Michael Elhadad. 2009. Hebrew dependency parsing: Initial results. In *Proceedings of IWPT*.

Yoav Goldberg and Reut Tsarfaty. 2008. A single framework for joint morphological segmentation and syntactic parsing. In *Proceedings of ACL*.

Yoav Goldberg, Meni Adler, and Michael Elhadad. 2006. Noun phrase chunking in hebrew: Influence of lexical and morphological features. In *Proceedings of COLING-ACL*.

Nomie Guthmann, Yuval Krymolowski, Adi Milea, and Yoad Winter. 2009. Automatic annotation of morpho-syntactic dependencies in a Modern Hebrew treebank. In Frank Van Eynde, Anette Frank, Koenraad De Smedt, and Gertjan van Noord, editors, *Proceedings of TLT*.

Kenneth L. Hale. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory*, 1(1).

Julia Hockenmaier and Mark Steedman. 2003. Parsing with generative models of predicate-argument structure. In *Proceedings of ACL*.

Dan Klein and Christopher D. Manning. 2003. Accurate unlexicalized parsing. In *Proceedings of ACL*.

Mohamed Maamouri, Ann Bies, and Seth Kulick. 2008. Enhanced annotation and parsing of the arabic treebank. In *Proceedings of INFOS*.

Yusuke Miyao and Jun'ichi Tsujii. 2008. Feature-forest models for probabilistic hpsg parsing. *Computational Linguistics*, 34(1):35–80.

Lilja Øvrelid and Joakim Nivre. 2007. Swedish dependency parsing with rich linguistic features. In *Proceeding of RANLP*.

Slav Petrov, Leon Barrett, Romain Thibaux, and Dan Klein. 2006. Learning accurate, compact, and interpretable tree annotation. In *Proceedings of ACL*.

Reut Tsarfaty and Khalil Sima'an. 2008. Relational-realizational parsing. In *Proceedings of CoLing*.

Reut Tsarfaty, Khalil Sima'an, and Remko Scha. 2009. An alternative to head-driven approaches for parsing a (relatively) free word order language. In *Proceedings of EMNLP*.