Argument Optionality in the LinGO Grammar Matrix

Safiyyah Saleem
University of Washington
ssaleem@u.washington.edu

Emily M. Bender
University of Washington
ebender@u.washington.edu

Abstract

We present a library of implemented HPSG analyses for argument optionality based on typological studies of this phenomenon in the world's languages, developed in the context of a grammar customization system that pairs a cross-linguistic core grammar with extensions for non-universal phenomena on the basis of user input of typological properties. Our analyses are compatible with multiple intersecting phenomena, including person, number, gender, tense, aspect and morphological rule formulation. We achieve 80-100% coverage on test suites from 10 natural languages.

1 Introduction

The LinGO Grammar Matrix customization system (Bender et al., 2002; 2010) is a web-based tool that creates starter grammars based on users' input to a questionnaire. The system comprises a core grammar covering linguistic phenomena that are posited to be universal (e.g. semantic compositionality) and a set of libraries providing analyses for phenomena that vary across languages (e.g. case). These resources are developed in the context of HPSG (Pollard and Sag, 1994), Minimal Recursion Semantics (Copestake et al., 2005), and the LKB grammar development environment (Copestake, 2002).

Previous to the work reported here, the Grammar Matrix customization system did not handle argument optionality—the possibility of leaving arguments unexpressed in lieu of overt pronouns. This phenomenon, also called pro-drop, argument drop, or null instantiation, is extremely common: according to Dryer (2008), 79% of the 674 languages sampled cannot or do not normally use independent pronouns in subject position. Accordingly, adding it to the customization system improves the system's ability to handle a large class of core sentences in many languages.

For example, in Modern Standard Arabic [arb] (Semitic), overt pronominal subjects are dropped in non-emphatic contexts. Previously, the system was able to model only the longer variant of (1).

(1) (hiyya)   naama-t
    (3.FEM.SG) sleep.PAST-3.FEM.SG
    She slept. [arb]

Furthermore, there was no way to adequately account for languages such as Hausa [hau] (Chadic) which do not allow overt simple pronominal subjects and prohibit overt objects after certain verb forms. The grammar would predict the opposite grammaticality for the examples in (2).

(2) (*nī)   nā-san   amsā
    (*1.SG) 1.SG.COMP-know answer
    I know the answer. [hau]

It might seem that these facts could be handled by adding a rule that allows arguments to be dropped if an appropriate option is checked in the customization system. However, the data from Arabic and Hausa suggest that such an approach would be insufficient, as languages place different constraints on the contexts in which overt arguments are required or prohibited.

In §2 we discuss the broad range of typological variation in argument optionality in the world's languages. In §3 we offer a set of HPSG analyses for these patterns. §4 explains how these analyses were incorporated into the Grammar Matrix

customization system and integrated with the existing libraries. We then present the results of a three-tiered evaluation of the implemented system in §5. The results demonstrate that the system is capable of accurately modeling the attested syntactic argument optionality patterns exhibited by a typologically diverse group of languages as well as the currently unattested but logically possible co-occurrence restrictions on affixes and overt arguments. To our knowledge, this is the first such system. The paper closes with a brief look at how the library could be extended even further to capture the range of semantic distinctions.

2 Typological Patterns

The typological literature shows that argument optionality is extremely common: Dryer (2008) found that of 674 geographically and genetically diverse languages, only 141 normally or obligatorily used independent pronominal subjects. Dryer distinguishes 4 categories in the remaining 533 languages, corresponding to how information about the person, number, and gender (PNG) of the subject is encoded: affixation on the verb, clitics on variable hosts, no encoding, or a mixed strategy. In addition, there are other dimensions in which languages vary, e.g., constraints on contexts in which dropping is done (see (1)–(2)).

Although we were unable to find a similar comprehensive survey of unexpressed objects, there is evidence to suggest that it too may be very widespread. In particular, lexically-licensed object dropping seems to be very common. Even English, which has a very strong preference for overt subjects, can be analyzed as licensing lexically-based object dropping (Fillmore, 1986). As with subject dropping, we also found a number of different co-occurrence restrictions on the presence of verbal affixes and overt objects. Some languages always encode the PNG of an object on the verb, others optionally do so if an overt object is present and obligatorily do so if one is not, while still others do not encode this information at all.

Drawing on work by Dryer and others, Table 1 summarizes the 6 major dimensions along which the rules licensing argument dropping differ. The first constraint is syntactic context. Most languages that license argument dropping do so re-

gardless of tense/aspect, mood, or person. Finnish [fin] and Hebrew [heb] are two notable exceptions (Vainikka and Levy, 1999).

The second constraint, lexically-based licensing, is most commonly found in object dropping. For example, while English usually prohibits argument dropping, it arguably licenses it with verbs such as 'found out', 'agree', and 'promise' (Fillmore, 1986). Lexically-based subject drop is found in Tamil [tam], which generally licenses subject dropping aside from some weather related verbs (Asher, 1985).

The third constraint, noun phrase type, captures the difference between a language such as Hausa which generally prohibits independent pronouns from appearing as subjects and other languages, which allow pronouns in subject position (possibly with emphatic interpretations).

The fourth constraint concerns the position of PNG markers. Of the languages with subject PNG markers and subject dropping, many encode subject PNG as a verbal affix. This pattern is exhibited by such geographically and genetically diverse languages as Spanish [spa], Arabic [arb], West Greenlandic [kal], Tamil [tam], and Nkore-Kiga [nyn]. Other languages such as Chemehuevi [ute], Polish [pol], and Warlpiri [wbp] make use of a clitic which can attach to different types of hosts (Dryer, 2008).

The final two constraints concern co-occurrence restrictions between PNG markers and overt objects. In some Bantu languages such as Nkore-Kiga, a verbal affix is not used unless the object precedes the verb or is pronominal. Object markers are not used when a full NP follows the verb (Taylor, 1985). In written French [fra], verbal affixes[1] are required if an object is dropped and not permitted if it is overt. In Arabic, for most transitive verbs, an object marker is required if an object is dropped and is optional if it is present. Hausa exhibits a more complex pattern: for tenses in which the verbal affix denoting PNG is morphologically separable from the tense marker, the PNG affix is optional if an overt noun phrase is present and required if it is not (Newman, 2000).

---

[1]See (Miller and Sag, 1997) for convincing arguments that so-called 'clitics' in French are actually affixes.

Constraint	(GF)	Possible Values
Syntactic context	(SUBJ)	{ All, select } tenses/aspects/moods/persons
Lexically-based	(SUBJ, OBJ)	{ All, select } verbs
Noun phrase type	(SUBJ, OBJ)	Independent pronouns { allowed, prohibited }
Placement of PNG marker	(SUBJ)	{ Verb, variable host }
PNG marking w/ dropped argument	(OBJ)	{ Required, optional, not permitted }
PNG marking w/ overt argument	(OBJ)	{ Required, optional, not permitted }

Table 1: Typological variation in licensing argument dropping

Noting these differences led us to posit that when an argument is dropped, there are three possibilities. A verbal affix can be: not permitted, optional, or required. The same three possibilities exist for overt objects as well. Combining what happens when an argument is dropped with what happens when it is present, gives us nine logically possible co-occurrence patterns.

Our review of the typological literature has shown that languages place different constraints on argument dropping. These constraints can be lexical, syntactic, or related to affixation and affix/overt-argument co-occurrence restrictions.

3 Analysis

This section presents HPSG analyses modeling the six dimensions of variation described in §2.

HPSG models natural language by positing lexical entries, lexical rules, and phrase structure rules, all described in terms of feature structures. A central idea, inspired by earlier work in Categorial Grammar (Ajdukiewicz, 1935; Bar-Hillel, 1953), is the notion of valence features. These list-valued features (including SUBJ and COMPS) contain information about the dependents required by a head. The valence lists are projected up the tree within the domain of each head, but shortened as the dependents are realized. A sentence is thus a verbal projection with empty SUBJ and COMPS lists.

In this context, argument dropping is the shortening of a valence list without the overt realization of the argument. Formally, this can be accomplished in at least three different ways: (1) In the mapping of arguments from the ARG-ST (argument structure) feature to the valence lists, one or more arguments can be suppressed, (2) lexical rules can operate on the valence lists, shortening them, or (3) unary (non-branching) phrase structure rules can cancel off valence elements. In this

work, we take the third approach, as we find it affords us the most flexibility to deal with variations across languages in constraints on argument optionality, while promoting similarity of analyses across languages.

We control the applicability of the unary-branching rules with the boolean feature OPT, marked on elements of valence lists.[2] For languages which allow subject/object dropping, we instantiate new phrase structure rules: *head-opt-subj-phrase* and/or *head-opt-comp-phrase*. These rules allow the head verb to satisfy a valence requirement without combining with another expression. To undergo these rules, the head daughter (the verb) must specify that the argument that is to be dropped is compatible with [OPT +]. This is sufficient to account for many languages. However, to ensure that languages which have lexical, syntactic context, and affix co-occurrence restrictions do not overgenerate, further additions to the grammar are necessary.

For lexical and affix-co-occurrence restrictions, we prevent overgeneration by manipulating the OPT feature. In languages which only license argument dropping for certain lexical items, we force those verbs which do not allow argument dropping to have arguments that are constrained to be [OPT −]. This prevents them from undergoing the subject/object dropping rules. Verbs are then classified into four different types based on whether or not they allow subject and/or object dropping. Individual lexical items instantiate these types. For those verbs which do not allow a particular argument to be dropped, the only way to satisfy the valence requirement is to combine with an overt argument.

---

[2]This feature was provided by the core Matrix but was not previously used in the customization system. To our knowledge it is not commonly used within HPSG analyses aside from in grammars that were derived from the Matrix.

Dropped/Overt Argument Affix	Overt Arg Rule	No-Marker-Rule	Marker-Rule	Transitive Verb Lex
required/required	underspecified	none	underspecified	needs lex rule
optional/optional	underspecified	none	underspecified	underspecified
not permitted/not permitted	underspecified	none	none	underspecified
required/optional	OPT −	OPT −	underspecified	needs lex rule
optional/not permitted	OPT −	none	OPT +	underspecified
not permitted/required	OPT −	OPT +	OPT −	needs lex rule
required/not permitted	OPT −	OPT −	OPT +	needs lex rule
optional/required	OPT −	OPT +	underspecified	needs lex rule
not permitted/optional	OPT −	none	OPT −	underspecified

Table 2: Constraints associated with logically possible affix co-occurrence

Languages with complex affix co-occurrence restrictions are modeled by manipulating the OPT feature in a different way: Constraints are placed on lexical and phrase structure rules, as well as on lexical types. In particular, we constrain the rules which combine verbs with overt arguments to check that that argument position is compatible with [OPT −]. This allows the lexical rules attaching the affixes to constrain the optionality of the corresponding argument position. In some of the nine logical possibilities, enforcing these constraints requires sending the verb through "no-marker" lexical rules so that constraints associated with markerless verbs can be enforced. Table 2 summarizes the constraints on the OPT feature on lexical and phrase structure rules, as well as the constraints on lexical types. The first column of this table lists the nine logically possible combinations described in §2. For example, the row labeled "required/required" gives the analysis for a language like West Greenlandic, which allows object dropping and always requires an object marker on the verb regardless of whether or not an overt object is present. In such a language, neither the lexical rules nor the overt-complement phrase structure rule constrain OPT, but the transitive verb lex type is required to undergo some object marking lexical rule.

For licensing that is based on syntactic context (subject dropping only) such as the Finnish and Hebrew examples presented in §2, we place constraints on the daughter of the unary subject drop rule which restrict its application to the right contexts. For example, to account for the argument optionality pattern present in Finnish, we constrain the *head-opt-subj-phrase* rule to require that the item on the head daughter's SUBJ list be spec-
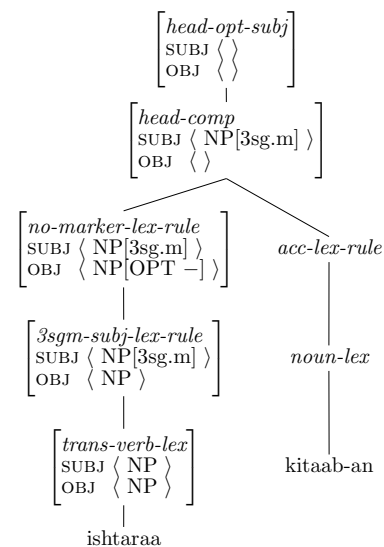

Figure 1: Parse structure for (3)

ified as non-third-person ([PER *non-third*]). Verbs not meeting this constraint are only allowed to empty their SUBJ lists by combining with an overt subject via the standard, binary *head-subj-phrase* rule. We have not seen a language which licenses subject dropping in syntactic contexts which do not form a natural class according to our feature system. However, our analysis easily lends itself to modeling this type of pattern if it exists by creating multiple different subtypes of the subject drop rule.

We close this section by illustrating our analysis with an example from Arabic. The sentence in (3) involves subject drop and an overt object. Since the object is overt, the verb bears only marking of subject PNG. The grammar that our system generates for Arabic assigns (3) the structure sketched in Figure 1.

(3) ishtaraa   kitaab-an
     3ms.buy.past book-acc
     He bought a book

4 Customized Grammar Creation

Before the addition of the argument optionality library, the phenomena covered in the Grammar Matrix customization system included word order, person, number, gender, case, tense/aspect, coordination, matrix yes-no questions, and sentential negation. The user is also allowed to specify lexical items and the morphological rules associated with each of them. Each of the phenomena correspond to a page of the questionnaire.

As the user answers questions, the choices are saved in a 'choices' file. The questionnaire is dynamic and the answers contained in the 'choices' file affect the types of features that the user is able to choose from on subsequent pages. For example, if the user describes the language as having 1st, 2nd, and 3rd persons on the Person page, then on the lexicon page, the user can create separate noun types for each person. Once the 'choices' file contains responses to required sections, the user is able to create the customized starter grammar by clicking on the 'create grammar' button. This invokes the customization script which uses the responses contained in the file to create a grammar that is compatible with the LKB grammar development environment.

Our implementation entailed additions to two major components of the system: the web-based questionnaire and the customization script. To determine which, if any, of the analyses presented in §3 should be included in the customized grammar, we needed to elicit the type of argument optionality pattern the language exhibited. Thus, we added an Argument Optionality page to the questionnaire. The page is divided into two sections—one for subject dropping and one for object dropping. In the section on subject dropping, the user is asked whether subject dropping exists and if so, whether it is context-dependent. For context-dependent subject dropping, the user is allowed to specify the syntactic contexts in which subject dropping is licensed by choosing from a multiselect list of features. There is the option to create multiple contexts. The features that appear in the list are drawn from those that the user chose on previous pages in the questionnaire. The user is also directed to select whether subject dropping

is lexically-based, whether affixes are required, optional or not permitted with overt arguments and whether affixes are required, optional or not permitted with dropped arguments. The questions presented in the object dropping section are identical to those in the subject dropping section with the exception that there is no question about context-dependent object dropping.

Since some of the constraints must be placed on individual lexical items and morphological rules, the page also includes instructions to the user on additional steps that need to be taken when completing the Lexicon page. For example, when describing a language where affixes are optional if an argument is dropped and not permitted if an overt argument is present, users are instructed to select 'overt-arg-not-permitted' for those affixes on the Lexicon page.

The changes to the customization script included adding each of the analyses described in §3 along with a mechanism for determining which of the analyses should be included in the grammar depending on the choices related to argument optionality, lexical items, and morphological rules contained in the 'choices' file. The resulting customized grammars include the rules and constraints necessary to allow and prohibit strings that do not contain overt arguments based upon the facts of a particular language as described by the user in the questionnaire.

5 Evaluation

The evaluation was conducted in a three stage process. Each stage involves constructing a set of test suites containing grammatical and ungrammatical strings representing the argument optionality pattern of a set of languages, generating grammars for the languages by answering the Grammar Matrix questionnaire, using the grammars to parse the sentences in the test suite, and hand-verifying the results. The three stages differed in the nature of the languages, the method by which the languages were selected, and the breadth of the customized grammars. The test suites are small, as they are specifically targeted at the phenomenon of argument optionality, but representative in the sense that they cover the space of relevant contrasts in each language.

5.1 Set 1: Pseudo-Languages

In the first stage, we tested the analyses presented in §3 by creating and then using the Grammar Matrix customization system to generate grammars for 38 pseudo-languages (sets of strings with associated grammaticality assignments) which collectively exhaustively exhibit each of the lexical, syntactic context or affix co-occurrence restriction patterns described in Table 1 (§2). All of the possible values identified for these given patterns are present in at least one language, as well as cross-classifications of different dimensions of constraints where appropriate. For example, there are pseudo-languages which share the property of always requiring object markers but differ in that one has lexically licensed object dropping and the other general object dropping. These pseudo-languages test the argument optionality analyses in isolation in that argument optionality is not constrained by other phenomena such as word order.

The customized grammars were able to accurately parse grammatical strings and rule out ungrammatical ones. Coverage on this set of 38 pseudo-languages was 100% with 0% overgeneration and no spurious ambiguity, thus validating the functioning of our analyses across the known typological space.

5.2 Set 2: Illustrative Languages

Next, we tested the system's performance in modeling part of a natural language. For this stage we deliberately chose several languages which exemplified interesting licensing and co-occurrence restriction patterns, including some which were considered during the development of the system. Each test suite included examples of grammatical and ungrammatical strings that were constructed based on the descriptions of the language given in the following sources: Suleiman 1990 (Arabic), Sulkala and Merja 1992 (Finnish), Newman 2000 (Hausa), and Asher 1985 (Tamil). As the test suites were designed to evaluate argument optionality, we restricted the test items to this phenomenon only. Other syntactic phenomena were only included if they affected the argument optionality pattern in the language. For example, gender distinctions were considered only for languages in which this was relevant to affix marking. A brief description of the argument optionality patterns found in these languages follows.

Arabic [arb] (Semitic) Pronominal subjects and objects are generally dropped. Subject affixes are always required whether or not an overt noun phrases is present. Affixes marking object person, number, and gender are required for strictly transitive verbs when an overt noun phrase is not present. Other transitive verbs appear to allow object drop without the object affix.

Finnish [fin] (Uralic) First and second person subjects are freely dropped and markers appear on the verb whether or not an overt noun phrase is present. Third person subjects are not allowed to be dropped with a referential interpretation; however, third person pronouns are obligatorily dropped for what Sulkala and Merja (1992) describe as a generic impersonal meaning. This description fits into what some linguists refer to as the fourth person—a non-referential impersonal syntactic/semantic distinction that is often realized in English as the impersonal pronoun *one*. Since Finnish shows evidence of further syntactic distinctions between generic and referential use of the third person marker, we have analyzed this marker as actually corresponding to two homophonous morphemes. One requires an overt subject and the other requires a dropped subject. There are no verbal affixes for PNG of the object.

Hausa [hau] (Chadic) Hausa generally requires pronominal subjects to be dropped. Simple, unmodified, uncoordinated independent pronouns are ungrammatical in subject position. Subject PNG is marked in a person aspect complex (PAC) along with tense and aspect information. The PAC precedes the lexical verb. When the PNG marker is morphologically segmentable from the tense/aspect, the PNG marker can be omitted if an overt noun phrase is present and is required if the noun phrase is not present. PNG is not marked for objects; however the verb form changes depending on whether a full noun phrase, pronoun, or no object immediately follows the verb.

Tamil [tam] (Dravidian) Subjects and objects can be freely dropped aside from a special class of weather verbs requiring overt subjects. Subject

PNG markers are always required whether a subject is overt or not. PNG is not marked for objects.

Lg.	Items	Grammatical	Ungrammatical	Coverage/ Overgeneration (%)
Arabic	13	10	3	90/0
Finnish	11	9	3	100/0
Hausa	20	8	12	100/0
Tamil	7	5	2	100/0

Table 3: Illustrative Languages Results

As shown in Table 3 we achieved 100% coverage over every test suite in this set except for Arabic. In addition, there was no overgeneration or spurious ambiguity. One Arabic item did not parse because the current implementation of our analyses does not elegantly account for obligatory object marking (with object drop) on some transitive verbs and optional object marking on others. We could have customized a grammar that included another, parallel set of lexical rules that would account for this item. Improvements to this aspect of the argument optionality library depend on upgrades to the morphotactic system.

5.3 Set 3: Held-out Languages

Finally, we tested a set of 'held out' languages not considered during development and chosen for their geographic and genetic diversity without regard for argument dropping patterns. We had previously created the non-argument optionality portions of these test suites and choices files to test the coverage of other libraries in the customization system and thus they include a wider variety of linguistic phenomena than Sets 1 or 2. As before, the construction and grammaticality judgements of the strings were based on descriptive grammars: Chirikba 2003 (Abkhaz), Press 1979 (Chemehuevi), Smirnova 1982 and Newman 2000 (Hausa), Pensalfini 2003 (Jingulu), Asher and Kumari 1997 (Malayalam), Taylor 1985 (Nkore-Kiga), and Fortescue 2003 (W. Greenlandic).

Due to space constraints, we provide only a summary of the argument optionality patterns in these languages (Table 4). All the languages licensed both subject and object dropping and in two of the six, dropping pronominal arguments was strongly preferred. Three languages have word order constraints on how argument option-

ality is realized: Abkhaz restricts the appearance of one of the third person affixes depending on verb-object order. Nkore-Kiga requires and prohibits the appearance of an object marker depending on where the overt object occurs. Chemehuevi requires that the clitic which is used to mark the subject appear in second position. It is also the only language that has lexical constraints on object dropping. Malayalam was the only language which did not mark person, number, and gender information for the subject.

The customized grammars were able to account for the majority of the patterns demonstrated in these languages (Table 5). We achieved 100% coverage on four languages with zero (Jingulu, Malayalam, West Greenlandic) or moderate (Abkhaz) overgeneration. The main source of errors found in the results is the handling of word order constraints: The grammars were unable to license (Chemehuevi) or restrict (Nkore-Kiga and Abkhaz) argument optionality based on the verb's and argument's positions in the sentence. Once the Grammar Matrix word order library has been improved and is able to account for second position clitics and fine-grained head-complement word order constraints, it will be a simple process to add the new feature(s) to existing lexical rules to account for these patterns. Incorporating the new functionality will not require any major changes to the argument optionality library aside from modifying the questionnaire to elicit the new information from the user.

Language	Items	Grammatical	Ungrammatical	Coverage/ Overgeneration (%)
Abkhaz	10	6	4	100/10
Chemehuevi	8	6	2	83.3/0
Jingulu	9	6	3	100/0
Malayalam	4	4	0	100/0
Nkore-Kiga	10	4	6	100/83.3
W. Greenlandic	5	3	2	100/0

Table 5: Held-out Language Results

In addition, we verified that the addition of argument optionality didn't reduce coverage on any other portion of these testsuites. This indicates that the new argument optionality library is interacting properly with existing libraries. Additional interactions will be tested as we add new libraries to the customization system.

	Object Dropping	Subject Dropping	Word Order Constraints	Lexical Constraints
Abkhaz	opt	opt	yes	none
Chemehuevi	opt	opt	yes	yes
Jingulu	opt	opt	none	none
Malayalam	opt	opt	no	none
Nkore-Kiga	pref	pref	yes	none
W. Greenlandic	pref	pref	none	none

Table 4: Existence of and constraints on argument optionality in six languages

6 Related Work

Subject dropping has been studied extensively within theoretical linguistics under many different frameworks (Rizzi, 1986; Bresnan, 2001; Ackema et al., 2006; Ginzburg and Sag, 2000). Within the context of HPSG, our analysis is similar to the one in the Grammar Matrix-derived Portuguese grammar (Branco and Costa, 2008) and to Müller's (2009) treatment of subject dropping in Maltese. These analyses differ from Ginzburg and Sag's (2000) HPSG analysis which uses language specific variations on the Argument Realization Principle to control whether the subject/object is placed onto the COMPS and/or SUBJ lists.

Language specific analyses have been implemented in deep, broad-coverage grammars for languages such as Japanese (Masuichi et al. (2003), Siegel and Bender (2002)) and Portuguese (Branco and Costa (2008)). Within the ParGram project (Butt et al., 2002), Kim et al. (2003) were able to directly port the argument optionality related rules from a Japanese grammar to Korean. However, to our knowledge, no one has implemented an analysis that has been applied to a large number of typologically, geographically, and genetically diverse languages.

7 Conclusion

Our current work has focused on modeling the variation in syntactic constraints on the licensing and restriction of argument dropping. To our knowledge, this is the first analysis of argument optionality that combines typological breadth with precision analyses that have been implemented and tested on a number of geographically and genetically diverse languages. Although we have tried to account for the patterns found in the typological literature, there may be variants that we are unaware of. We hope to learn of more patterns as the Grammar Matrix customization system is applied to an ever wider set languages.

While the current work focuses on syntactic variation, we intend to expand the argument optionality library to include semantic distinctions as well. A likely starting point would be the proposal given by Bender and Goss-Grubbs (2008) who present a way to model the discourse status (Prince, 1981) of an NP taking into account the differences between definite and indefinite null instantiation described by Fillmore (1986). In addition, ongoing work to improve the word order library may eventually allow us to more accurately model word-order based constraints.

References

Ackema, Peter, Patrick Brandt, Maaike Schoorlemmer, and Fred Weerman, editors. 2006. *Arguments and Agreement*. Oxford University Press, Oxford.

Ajdukiewicz, Kazimierz. 1935. Die syntaktische konnexität. *Studia Philosophica*, 1:1–27.

Asher, R.E. and T.C. Kumari. 1997. *Malayalam*. Routledge, NY.

Asher, R.E. 1985. *Tamil*. Croom Helm, London.

Bar-Hillel, Yehoshua. 1953. A quasi-arithmetical notation for syntactic description. *Language*, 29:47–58.

Bender, Emily M. and David Goss-Grubbs. 2008. Semantic representations of syntactically marked discourse status in crosslinguistic perspective. In *Proc. 2008 Conference on Semantics in Text Processing*, pages 17–29.

Bender, Emily M., Dan Flickinger, and Stephan Oepen. 2002. The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In *Proc. Workshop on Grammar Engineering and Evaluation at COLING 2002*, pages 8–14.

Bender, Emily M., Scott Drellishak, Antske Fokkens, Michael Wayne Goodman, Daniel P. Mills, Laurie Poulson, and Safiyyah Saleem. 2010. Grammar prototyping and testing with the LinGO Grammar Matrix customization system. In *Proc. ACL 2010 Software Demonstrations*.

Branco, António and Francisco Costa. 2008. A computational grammar for deep linguistic processing of Portuguese: LXGram, version a.4.1. Technical report, University of Lisbon, Dept. of Informatics.

Bresnan, Joan. 2001. *Lexical Functional Syntax*. Blackwell, Boston.

Butt, Miriam, Helge Dyvik, Tracy Holloway King, Hiroshi Masuichi, and Christian Rohrer. 2002. The parallel grammar project. In *Proc. Workshop on Grammar Engineering and Evaluation at COLING 2002*, pages 1–7.

Chirikba, Viachesiav. 2003. *Abkhaz*. LINCOM, Munich.

Copestake, Ann, Dan Flickinger, Carl Pollard, and Ivan A. Sag. 2005. Minimal recursion semantics: An introduction. *Research on Language & Computation*, 3(4):281–332.

Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars*. CSLI, Stanford.

Dryer, Matthew. 2008. Expression of proniminal subjects. In Haspelmath, Martin, Matthew Dryer, David Gil, and Bernard Comrie, editors, *The World Atlas of Language Structures Online*, chapter 101. Max Planck Digital Library.

Fillmore, Charles. 1986. Pragmatically controlled zero anaphora. In *Proc. 12th annual meeting of the Berkeley Linguistics Society*, pages 95–107.

Fortescue, Michael. 2003. *West Greenlandic*. Croom Helm, London.

Ginzburg, Johnathan and Ivan Sag. 2000. *Interrogative Investigations*. CSLI, Stanford.

Kim, Roger, Mary Dalrymple, Ronald M. Kaplan, Tracy Holloway King, Hiroshi Masuichi, and Tomoko Ohkuma. 2003. Multilingual grammar development via grammar porting. In *ESSLLI 2003 Workshop on Ideas and Strategies for Multilingual Grammar Development*, pages 49–56.

Masuichi, Hiroshi, Tomoko Ohkuma, Hiroki Yoshimura, and Yasunari Harada. 2003. Japanese parser on the basis of the lexical-functional grammar formalism and its evaluation. In Dong Hong Ji, Kim Teng Lua, editor, *Proc. PACLIC17*, pages 298–309.

Miller, Philip H. and Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language & Linguistic Theory*, 15(3):573–639.

Müller, Stefan. 2009. Towards an HPSG analysis of Maltese. In et al, Bernard Comrie, editor, *Introducing Maltese linguistics. Papers from the 1st International Conference on Maltese Linguistics*, pages 83–112. Benjamins, Amsterdam.

Newman, Paul. 2000. *The Hausa Language: An encyclopedic reference grammar*. Yale University Press, New Haven.

Pensalfini, Rob. 2003. *A Grammar of Jingulu: An Aboriginal language of the Northern Territory*. Pacific Linguistics, Canberra.

Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago, IL.

Press, Margaret. 1979. *Chemehuevi: A grammar and lexicon*. University of California Press, Berkeley.

Prince, Ellen. 1981. Toward a taxonomy of given-new information. In Cole, P., editor, *Radical Pragmatics*, pages 223–255. Academic Press, NY.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry*, 17(3):501–557.

Siegel, Melanie and Emily M. Bender. 2002. Efficient deep processing of Japanese. In *Proc. 3rd Workshop on Asian Language Resources and International Standardization at COLING 2002*.

Smirnova, Mirra A. 1982. *The Hausa Language: A Descriptive Grammar*. Routledge, Boston.

Suleiman, Saleh M. 1990. The semantic functions of object deletion in classical arabic. *Language Sciences*, 12(2-3):255 – 266.

Sulkala, Helena and Karjalaninen Merja. 1992. *Finnish*. Routledge, NY.

Taylor, Charles. 1985. *Nkore-Kiga*. Croom Helm, London.

Vainikka, Anne and Yonata Levy. 1999. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory*, 17:613–671.