Evaluating a Crosslinguistic Grammar Resource:
A Case Study of Wambaya

Emily M. Bender

University of Washington

Department of Linguistics

Box 354340

Seattle WA 98195-4340

`ebender@u.washington.edu`

Abstract

This paper evaluates the LinGO Grammar Matrix, a cross-linguistic resource for the development of precision broad coverage grammars, by applying it to the Australian language Wambaya. Despite large typological differences between Wambaya and the languages on which the development of the resource was based, the Grammar Matrix is found to provide a significant jump-start in the creation of the grammar for Wambaya: With less than 5.5 person-weeks of development, the Wambaya grammar was able to assign correct semantic representations to 76% of the sentences in a naturally occurring text. While the work on Wambaya identified some areas of refinement for the Grammar Matrix, 59% of the Matrix-provided types were invoked in the final Wambaya grammar, and only 4% of the Matrix-provided types required modification.

1 Introduction

Hand-built grammars are often dismissed as too expensive to build on the one hand, and too brittle on the other. Nevertheless, they are key to various NLP applications, including those benefiting from deep natural language understanding (e.g., textual inference (Bobrow et al., 2007)), generation of well-formed output (e.g., natural language weather alert systems (Lareau and Wanner, 2007)) or both (as in machine translation (Oepen et al., 2007)). Of particular interest here are applications concerning endangered languages: Endangered languages represent a case of minimal linguistic resources, typically lacking even moderately-sized corpora, let alone

treebanks. In the best case, one finds well-crafted descriptive grammars, bilingual dictionaries, and a handful of translated texts. The methods of precision grammar engineering are well-suited to taking advantage of such resources. At the same time, the applications of interest in the context of endangered languages emphasize linguistic precision: implemented grammars can be used to enrich existing linguistic documentation, to build grammar checkers in the context of language standardization, and to create software language tutors in the context of language preservation efforts.

The LinGO Grammar Matrix (Bender et al., 2002; Bender and Flickinger, 2005; Drellishak and Bender, 2005) is a toolkit for reducing the cost of creating broad-coverage precision grammars by prepackaging both a cross-linguistic core grammar and a series of libraries of analyses of cross-linguistically variable phenomena, such as major-constituent word order or question formation. The Grammar Matrix was developed initially on the basis of broad-coverage grammars for English (Flickinger, 2000) and Japanese (Siegel and Bender, 2002), and has since been extended and refined as it has been used in the development of broad-coverage grammars for Norwegian (Hellan and Haugereid, 2003), Modern Greek (Kordoni and Neu, 2005), and Spanish (Marimon et al., 2007), as well as being applied to 42 other languages from a variety of language families in a classroom context (Bender, 2007).

This paper aims to evaluate both the utility of the Grammar Matrix in jump-starting precision grammar development and the current state of its cross-linguistic hypotheses through a case study of a

language typologically very different from any of the languages above: the non-Pama-Nyungan Australian language Wambaya (Nordlinger, 1998).

The remainder of this paper is structured as follows: §2 provides background on the Grammar Matrix and Wambaya, and situates the project with respect to related work. §3 presents the implemented grammar of Wambaya, describes its development, and evaluates it against unseen, naturally occurring text. §4 uses the Wambaya grammar and its development as one point of reference to measure the usefulness and cross-linguistic validity of the Grammar Matrix. §5 provides further discussion.

2 Background

2.1 The LinGO Grammar Matrix

The LinGO Grammar Matrix is situated theoretically within Head-Driven Phrase Structure Grammar (HPSG; Pollard and Sag, 1994), a lexicalist, constraint-based framework. Grammars in HPSG are expressed as a collection of typed feature structures which are arranged into a hierarchy such that information shared across multiple lexical entries or construction types is represented only on a single supertype. The Matrix is written in the TDL (type description language) formalism, which is interpreted by the LKB parser, generator, and grammar development environment (Copestake, 2002). It is compatible with the broader range of DELPH-IN tools, e.g., for machine translation (Lønning and Oepen, 2006), treebanking (Oepen et al., 2004) and parse selection (Toutanova et al., 2005).

The Grammar Matrix consists of a cross-linguistic core type hierarchy and a collection of phenomenon-specific libraries. The core type hierarchy defines the basic feature geometry, the ways that heads combine with arguments and adjuncts, linking types for relating syntactic to semantic arguments, and the constraints required to compositionally build up semantic representations in the format of Minimal Recursion Semantics (Copestake et al., 2005; Flickinger and Bender, 2003). The libraries provide collections of analyses for cross-linguistically variable phenomena. The current libraries include analyses of major constituent word order (SOV, SVO, etc), sentential negation, coordination, and yes-no question formation. The Matrix is accessed through

a web-based configuration system[1] which elicits typological information from the user-linguist through a questionnaire and then outputs a grammar consisting of the Matrix core plus selected types and constraints from the libraries according to the specifications in the questionnaire.

2.2 Wambaya

Wambaya is a recently extinct language of the West Barkly family from the Northern Territory in Australia (Nordlinger, 1998). Wambaya was selected for this project because of its typological properties and because it is extraordinarily well-documented by Nordlinger in her 1998 descriptive grammar.

Perhaps the most striking feature of Wambaya is its word order: it is a radically non-configurational language with a second position auxiliary/clitic cluster. That is, aside from the constraint that verbal clauses require a clitic cluster (marking subject and object agreement and tense, aspect and mood) in second position, the word order is otherwise free, to the point that noun phrases can be non-contiguous, with head nouns and their modifiers separated by unrelated words. Furthermore, head nouns are generally not required: argument positions can be instantiated by modifiers only, or, if the referent is clear from the context, by no nominal constituent of any kind. It has a rich system of case marking, and adnominal modifiers agree with the heads they modify in case, number, and four genders. An example is given in (1) (Nordlinger, 1998, 223).[2]

(1) Ngaragana-nguja ngiy-a
     grog-PROP.IV.ACC 3.SG.NM.A-PST
     gujinganjanga-ni jiyawu ngabulu.
     mother.II.ERG   give   milk.IV.ACC
     '(His) mother gave (him) milk with grog in it.' [wmb]

In (1), *ngaragana-nguja* ('grog-proprietive', or 'having grog') is a modifier of *ngabulu* milk. They agree in case (accusative) and gender (class IV), but they are not contiguous within the sentence.

To relate such discontinuous noun phrases to appropriate semantic representations where 'having-

---

grog' and 'milk' are predicated of the same entity requires a departure from the ordinary way that heads are combined with arguments and modifiers combined with heads in HPSG in general and in the Matrix in particular.[3] In the Grammar Matrix, as in most work in HPSG, lexical heads record the dependents they require in valence lists (SUBJ, COMPS, SPR). When a head combines with one of its arguments, the result is a phrase with the same valence requirements as the head daughter, minus the one corresponding to the argument that was just satisfied. In contrast, the project described here has explored a non-cancellation analysis for Wambaya: even after a head combines with one of its arguments, that argument remains on the appropriate valence list of the mother, so that it is visible for further combination with modifiers. In addition, heads can combine directly with modifiers of their arguments (as opposed to just modifiers of themselves).

Argument realization and the combination of heads and modifiers are fairly fundamental aspects of the system implemented in the Matrix. In light of the departure described above, it is interesting to see to what extent the Matrix can still support rapid development of a precision grammar for Wambaya.

2.3 Related Work

There are currently many multilingual grammar engineering projects under active development, including ParGram, (Butt et al., 2002; King et al., 2005), the MetaGrammar project (Kinyon et al., 2006), KPML (Bateman et al., 2005), Grammix (Müller, 2007) and OpenCCG (Baldridge et al., 2007). Among approaches to multilingual grammar engineering, the Grammar Matrix's distinguishing characteristics include the deployment of a shared core grammar for crosslinguistically consistent constraints and a series of libraries modeling varying linguistic properties. Thus while other work has successfully exploited grammar porting between typologically related languages (e.g., Kim et al., 2003), to my knowledge, no other grammar porting project has covered the same typological dis-

tance attempted here. The current project is also situated within a broader trend of using computational linguistics in the service of endangered language documentation (e.g., Robinson et al., 2007, see also www.emeld.org).

3 Wambaya grammar

3.1 Development

The Wambaya grammar was developed on the basis of the grammatical description in Nordlinger 1998, including the Wambaya-English translation lexicon and glosses of individual example sentences. The development test suite consisted of all 794 distinct positive examples from Ch. 3–8 of the descriptive grammar. This includes elicited examples as well as (sometimes simplified) naturally occurring examples. They range in length from one to thirteen words (mean: 3.65). The test suite was extracted from the descriptive grammar at the beginning of the project and used throughout with only minor refinements as errors in formatting were discovered. The regression testing facilities of [incr tsdb()] allowed for rapid experimentation with alternative analyses as new phenomena were brought into the grammar (cf. Oepen et al., 2002).

With no prior knowledge of this language beyond its most general typological properties, we were able to develop in under 5.5 person-weeks of development time (210 hours) a grammar able to assign appropriate analyses to 91% of the examples in the development set.[4] The 210 hours include 25 hours of an RA's time entering lexical entries, 7 hours spent preparing the development test suite, and 15 hours treebanking (using the LinGO Redwoods software (Oepen et al., 2004) to annotate the intended parse for each item). The remainder of the time was ordinary grammar development work.[5]

In addition, this grammar has relatively low ambiguity, assigning on average 11.89 parses per item in the development set. This reflects the fact that the grammar is modeling grammaticality: the rules are

---

[3]A linearization-based analysis as suggested by Donohue and Sag (1999) for discontinuous constituents in Warlpiri (another Australian language), is not available, because it relies on disassociating the constituent structure from the surface order of words in a way that is not compatible with the TDL formalism.

[4]An additional 6% received some analysis, but not one that matched the translation given in the reference grammar.

[5]These numbers do not include the time put into the original field work and descriptive grammar work. Nordlinger (p.c.) estimates that as roughly 28 linguist-months, plus the native speaker consultants' time.

meant to exclude ungrammatical strings as well as are unwarranted analyses of grammatical strings.

3.2 Scope

The grammar encodes mutually interoperable analyses of a wide variety of linguistic phenomena, including:

• Word order: second position clitic cluster, otherwise free word order, discontinuous noun phrases

• Argument optionality: argument positions with no overt head

• Linking of syntactic to semantic arguments

• Case: case assignment by verbs to dependents

• Agreement: subject and object agreement in person and number (and to some extent gender) marked in the clitic cluster, agreement between nouns and adnominal modifiers in case, number and gender

• Lexical adverbs, including manner, time, and location, and adverbs of negation, which vary by clause type (declarative, imperative, or interrogative)

• Derived event modifiers: nominals (nouns, adjectives, noun phrases) used as event modifiers with meaning dependent on their case marking

• Lexical adjectives, including demonstratives adverbs, numerals, and possessive adjectives, as well as ordinary intersective adjectives

• Derived nominal modifiers: modifiers of nouns derived from nouns, adjectives and verbs, including the proprietive, privative, and 'origin' constructions

• Subordinate clauses: clausal complements of verbs like "tell" and "remember", non-finite subordinate clauses such as purposives ("in order to") and clauses expressing prior or simultaneous events

• Verbless clauses: nouns, adjectives, and adverbs, lexical or derived, functioning as predicates

• Illocutionary force: imperatives, declaratives, and interrogatives (including *wh* questions)

• Coordination: of clauses and noun phrases

• Other: inalienable possession, secondary predicates, causatives of verbs and adjectives

3.3 Sample Analysis

This section provides a brief description of the analysis of radical non-configurationality in order to give a sense of the linguistic detail encoded in the Wambaya grammar and give context for the evaluation of the Wambaya grammar and the Grammar Matrix in later sections.

The linguistic analyses encoded in the grammar serve to map the surface strings to semantic representations (in Minimal Recursion Semantics (MRS) format (Copestake et al., 2005)). The MRS in Figure 1 is assigned to the example in (1).[6] It includes the basic propositional structure: a situation of 'giving' in which the first argument, or agent, is 'mother', the second (recipient) is some third-person entity, and the third (patient), is 'milk' which is also related to 'grog' through the proprietive relation. It is marked as past tense, and as potentially a statement or a question, depending on the intonation.[7,8]

A simple tree display of the parse giving rise to this MRS is given in Figure 2. The non-branching nodes at the bottom of the tree represent the lexical rules which associate morphosyntactic information with a word according to its suffixes. The general left-branching structure of the tree is a result of the analysis of the second-position clitic cluster: The clitic clusters are treated as argument-composition auxiliaries, which combine with a lexical verb and 'inherit' all of the verb's arguments. The auxiliaries first pick up all dependents to the right, and then combine with exactly one constituent to the left.

The grammar is able to connect x7 (the index of 'milk') to both the ARG3 position of the 'give' relation and the ARG1 position of the proprietive relation, despite the separation between *ngaraganaguja* ('grog-PROP.IV.ACC') and *ngabulu* ('milk.IV.ACC') in the surface structure, as follows: The auxiliary *ngiya* is subject to the constraints in (2), meaning that it combines with a verb as its first complement and then the verb's complements as its remaining complements.[9] The auxiliary can combine with its complements in any order, thanks to a series of head-complement rules which realize the *n*th element of

---

[6]The grammar in fact finds 42 parses for this example. The one associated with the MRS in Figure 1 best matches the intended interpretation as indicated by the gloss of the example.

[7]The relations are given English predicate names for the convenience of the grammar developer, and these are not intended as any kind of interlingua.

[8]This MRS is 'fragmented' in the sense that the labels of several of the elementary predications (eps) are not related to any argument position of any other ep. This is related to the fact that the grammar doesn't yet introduce quantifiers for any of the nominal arguments.

[9]In this and other attribute value matrices displayed, feature paths are abbreviated and detail not relevant to the current point is suppressed.

$$\begin{bmatrix} \text{LTOP} & \text{h1} \\ \text{INDEX} & \text{e2 (prop-or-ques, past)} \\[4pt] \text{RELS} & \left\langle \begin{bmatrix} \mathbf{\_grog\_n\_rel} \\ \text{LBL} \quad \text{h3} \\ \text{ARG0} \quad \text{x4 (3, iv)} \end{bmatrix}, \begin{bmatrix} \mathbf{proprietive\_a\_rel} \\ \text{LBL} \quad \text{h5} \\ \text{ARG0} \quad \text{e6} \\ \text{ARG1} \quad \text{x7 (3, iv)} \\ \text{ARG2} \quad \text{x4} \end{bmatrix}, \begin{bmatrix} \mathbf{\_mother\_n\_rel} \\ \text{LBL} \quad \text{h8} \\ \text{ARG0} \quad \text{x9 (3sg, ii)} \end{bmatrix}, \begin{bmatrix} \mathbf{\_give\_v\_rel} \\ \text{LBL} \quad \text{h1} \\ \text{ARG0} \quad \text{e2} \\ \text{ARG1} \quad \text{x9} \\ \text{ARG2} \quad \text{x10 (3)} \\ \text{ARG3} \quad \text{x7} \end{bmatrix}, \begin{bmatrix} \mathbf{\_milk\_n\_rel} \\ \text{LBL} \quad \text{h5} \\ \text{ARG0} \quad \text{x7} \end{bmatrix} \right\rangle \\[4pt] \text{HCONS} & \langle \, \rangle \end{bmatrix}$$
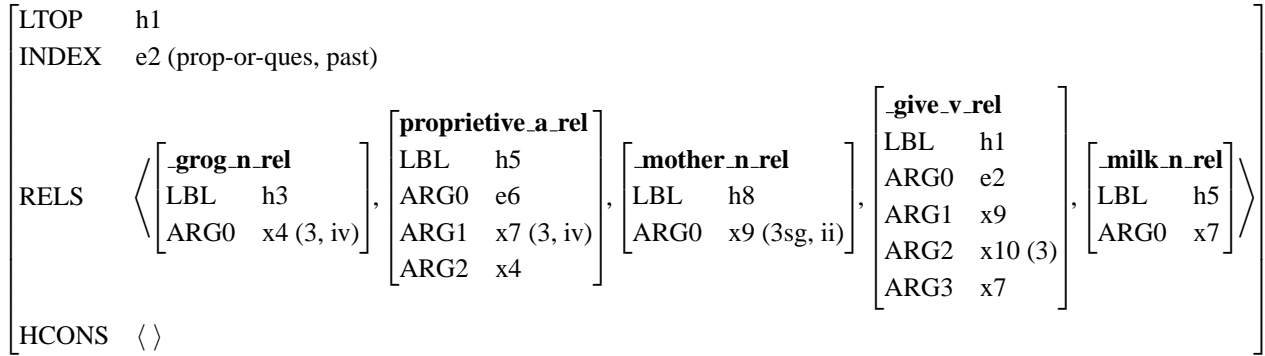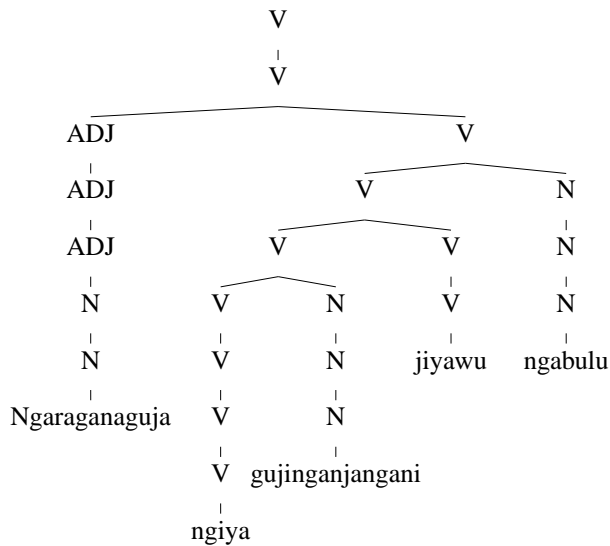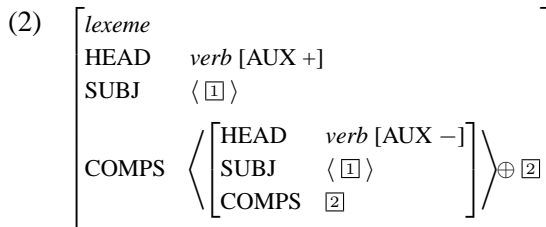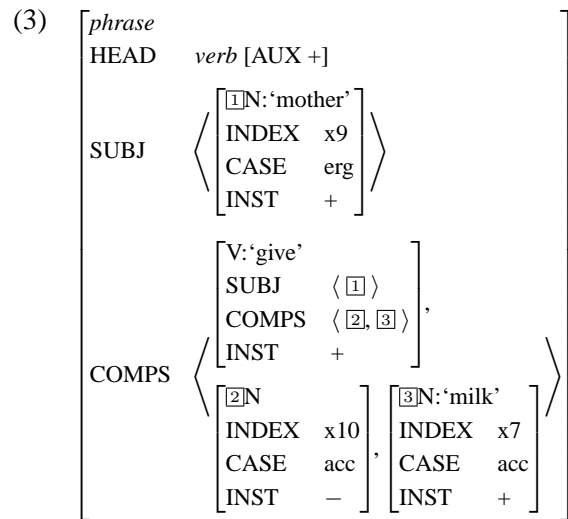
Figure 1: MRS for (1)

Figure 2: Phrase structure tree for (1)

the COMPS list. It this example, it first picks up the subject *gujinganjangani* ('mother-ERG'), then the main verb *jiyawu* ('give'), and then the object *ngabulu* ('milk-ACC').

$$(2) \quad \begin{bmatrix} lexeme \\ \text{HEAD} \quad verb\ [\text{AUX} +] \\ \text{SUBJ} \quad \langle \boxed{1} \rangle \\[4pt] \text{COMPS} \quad \left\langle \begin{bmatrix} \text{HEAD} & verb\ [\text{AUX} -] \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \boxed{2} \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix}$$

The resulting V node over *ngiya gujinganjangani jiyawu ngabulu* is associated with the constraints sketched in (3).

$$(3) \quad \begin{bmatrix} phrase \\ \text{HEAD} \quad verb\ [\text{AUX} +] \\[4pt] \text{SUBJ} \quad \left\langle \begin{bmatrix} \boxed{1}\text{N:'mother'} \\ \text{INDEX} & \text{x9} \\ \text{CASE} & \text{erg} \\ \text{INST} & + \end{bmatrix} \right\rangle \\[4pt] \text{COMPS} \quad \left\langle \begin{bmatrix} \text{V:'give'} \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2}, \boxed{3} \rangle \\ \text{INST} & + \end{bmatrix}, \begin{bmatrix} \boxed{2}\text{N} \\ \text{INDEX} & \text{x10} \\ \text{CASE} & \text{acc} \\ \text{INST} & - \end{bmatrix}, \begin{bmatrix} \boxed{3}\text{N:'milk'} \\ \text{INDEX} & \text{x7} \\ \text{CASE} & \text{acc} \\ \text{INST} & + \end{bmatrix} \right\rangle \end{bmatrix}$$

Unlike in typical HPSG approaches, the information about the realized arguments is still exposed in the COMPS and SUBJ lists of this constituent.[10] This makes the necessary information available to separately-attaching modifiers (such as *ngaraganaguja* ('grog-PROP.IV.ACC')) so that they can check for case and number/gender compatibility and connect the semantic index of the argument they modify to a role in their own semantic contribution (in this case, the ARG1 of the 'proprietive' relation).

3.4 Evaluation

The grammar was evaluated against a sample of naturally occurring data taken from one of the texts transcribed and translated by Nordlinger (1998) ("The two Eaglehawks", told by Molly Nurlanyma Grueman). Of the 92 sentences in this text, 20 overlapped with items in the development set, so the

---

[10]The feature INST, newly proposed for this analysis, records the fact that they have been instantiated by lexical heads.

	correct	parsed incorrect	unparsed	average ambiguity
Existing vocab	50%	8%	42%	10.62
w/added vocab	**76%**	8%	14%	12.56

Table 1: Grammar performance on held-out data

evaluation was carried out only on the remaining 72 sentences. The evaluation was run twice: once with the grammar exactly as is, including the existing lexicon, and a second time after new lexical entries were added, using only existing lexical types. In some cases, the orthographic components of the lexical rules were also adjusted to accommodate the new lexical entries. In both test runs, the analyses of each test item were hand-checked against the translation provided by Nordlinger (1998). An item is counted as correctly analyzed if the set of analyses returned by the parser includes at least one with an MRS that matches the dependency structure, illocutionary force, tense, aspect, mood, person, number, and gender information indicated.

The results are shown in Table 1: With only lexical additions, the grammar was able to assign a correct parse to 55 (76%) of the test sentences, with an average ambiguity over these sentences of 12.56 parses/item.

3.5 Parse selection

The parsed portion of the development set (732 items) constitutes a sufficiently large corpus to train a parse selection model using the Redwoods disambiguation technology (Toutanova et al., 2005). As part of the grammar development process, the parses were annotated using the Redwoods parse selection tool (Oepen et al., 2004). The resulting treebank was used to select appropriate parameters by 10-fold cross-validation, applying the experimentation environment and feature templates of (Velldal, 2007). The optimal feature set included 2-level grandparenting, 3-grams of lexical entry types, and both constituent weight features. In the cross-validation trials on the development set, this model achieved a parse selection accuracy of 80.2% (random choice baseline: 23.9%). A model with the same features was then trained on all 544 ambiguous examples

from the development set and used to rank the parses of the test set. It ranked the correct parse (exact match) highest in 75.0% of the test sentences. This is well above the random-choice baseline of 18.4%, and affirms the cross-linguistic validity of the parse-selection techniques.

3.6 Summary

This section has presented the Matrix-derived grammar of Wambaya, illustrating its semantic representations and analyses and measuring its performance against held-out data. I hope to have shown the grammar to be reasonably substantial, and thus an interesting case study with which to evaluate the Grammar Matrix itself.

4 Evaluation of Grammar Matrix

It is not possible to directly compare the development of a grammar for the same language, by the same grammar engineer, with and without the assistance of the Grammar Matrix. Therefore, in this section, I evaluate the usefulness of the Grammar Matrix by measuring the extent to which the Wambaya grammar as developed makes use of types defined in Matrix as well as the extent to which Matrix-defined types had to be modified. The former is in some sense a measure of the usefulness of the Matrix, and the latter is a measure of its correctness.

While the libraries and customization system were used in the initial grammar development, this evaluation primarily concerns itself with the Matrix core type hierarchy. The customization-provided Wambaya-specific type definitions for word order, lexical types, and coordination constructions were used for inspiration, but most needed fairly extensive modification. This is particularly unsurprising for basic word order, where the closest available option ("free word order") was taken, in the absence of a pre-packaged analysis of non-configurationality and second-position phenomena. The other changes to the library output were largely side-effects of this fundamental difference.

Table 2 presents some measurements of the overall size of the Wambaya grammar. Since HPSG grammars consist of types organized into a hierarchy and instances of those types, the unit of measure for these evaluations will be types and/or instances. The

		N
Matrix types		891
ordinary	390	
pos disjunctions	591	
Wambaya-specific types		911
Phrase structure rules		83
Lexical rules		161
Lexical entries		1528

Table 2: Size of Wambaya grammar

	Matrix core types		w/ POS types	
Directly used	132	34%	136	15%
Indirectly used	98	25%	584	66%
Total types used	230	59%	720	81%
Types unused	160	41%	171	19%
Types modified	16	4%	16	2%
Total	390	100%	891	100%

Table 3: Matrix core types used in Wambaya grammar

Wambaya grammar includes 891 types defined in the Matrix core type hierarchy. These in turn include 390 ordinary types, and 591 'disjunctive' types, the powerset of 9 part of speech types. These are provided in the Matrix so that Matrix users can easily refer to classes of, say, "nouns and verbs" or "nouns and verbs and adjectives". The Wambaya-specific portion of the grammar includes 911 types. These types are invoked in the definitions of the phrase structure rules, lexical rules, and lexical entries.

Including the disjunctive part-of-speech types, just under half (49%) of the types in the grammar are provided by the Matrix. However, it is necessary to look more closely; just because a type is provided in the Matrix core hierarchy doesn't mean that it is invoked by any rules or lexical entries of the Wambaya grammar. The breakdown of types used is given in Table 3. Types that are used directly are either called as supertypes for types defined in the Wambaya-specific portion of the grammar, or used as the value of some feature in a type constraint in the Wambaya-specific portion of the grammar. Types that are used indirectly are either ancestor types to types that are used directly, or types that are used as the value of a feature in a constraint in the Matrix core types on a type that is used (directly or indirectly) by the Wambaya-specific portion of the grammar.

Relatively few (16) of the Matrix-provided types needed to be modified. These were types that were useful, but somehow unsuitable, and typically deeply interwoven into the type system, such that not using and them and defining parallel types in their place would be inconvenient.

Setting aside the types for part of speech disjunctions, 59% of the Matrix-provided types are invoked by the Wambaya-specific portion of the grammar. While further development of the Wambaya grammar might make use of some of the remaining 41% of the types, this work suggests that there is a substantial amount of information in the Matrix core type hierarchy which would better be stored as part of the typological libraries. In particular, the analyses of argument realization implemented in the Matrix were not used for this grammar. The types associated with argument realization in configurational languages should be moved into the word-order library, which should also be extended to include an analysis of Wambaya-style radical non-configurationality. At the same time, the lexical amalgamation analysis of the features used in long-distance dependencies (Sag, 1997) was found to be incompatible with the approach to argument realization in Wambaya, and a phrasal amalgamation analysis was implemented instead. This again suggests that lexical v. phrasal amalgamation should be encoded in the libraries, and selected according to the word order pattern of the language.

As for parts of speech, of the nine types provided by the Matrix, five were used in the Wambaya grammar (verb, noun, adj, adv, and det) and four were not (num, conj, comp, and adp(osition)). Four disjunctive types were directly invoked, to describe phenomena applying to nouns and adjectives, verbs and adverbs, anything but nouns, and anything but determiners. While it was convenient to have the disjunctive types predefined, it also seems that a much smaller set of types would suffice in this case. Since the nine proposed part of speech types have varying crosslinguistic validity (e.g., not all languages have conjunctions), it might be better to provide software support for creating the disjunctive types as the need arises, rather than predefining them.

Even though the number of Matrix-provided types is small compared to the grammar as a whole, the relatively short development time indicates that the types that were incorporated were quite useful. In providing the fundamental organization of the gram-

mar, to the extent that that organization is consistent with the language modeled, these types significantly ease the path to creating a working grammar.

The short development time required to create the Wambaya grammar presents a qualitative evaluation of the Grammar Matrix as a crosslinguistic resource, as one goal of the Grammar Matrix is to reduce the cost of developing precision grammars. The fact that a grammar capable of assigning valid analyses to an interesting portion of sentences from naturally occurring text could be developed in less than 5.5 person-weeks of effort suggests that this goal is indeed met. This is particularly encouraging in the case of endangered and other resource-poor languages. A grammar such as the one described here could be a significant aide in analyzing additional texts as they are collected, and in identifying constructions that have not yet been analyzed (cf. Baldwin et al, 2005).

5 Conclusion

This paper has presented a precision, hand-built grammar for the Australian language Wambaya, and through that grammar a case study evaluation of the LinGO Grammar Matrix. True validation of the Matrix *qua* hypothesized linguistic universals requires many more such case studies, but this first test is promising. Even though Wambaya is in some respects very different from the well-studied languages on which the Matrix is based, the existing machinery otherwise worked quite well, providing a significant jump-start to the grammar development process. While the Wambaya grammar has a long way to go to reach the complexity and range of linguistic phenomena handled by, for example, the LinGO English Resource Grammar, it was shown to provide analyses of an interesting portion of a naturally occurring text. This suggests that the methodology of building such grammars could be profitably incorporated into language documentation efforts.

The Grammar Matrix allows new grammars to directly leverage the expertise in grammar engineering gained in extensive work on previous grammars of better-studied languages. Furthermore, the design of the Matrix is such that it is not a static object, but intended to evolve and be refined as more languages are brought into its purview. Generalizing the core hierarchy and libraries of the Matrix to support languages like Wambaya can extend its typological reach and further its development as an investigation in computational linguistic typology.

References

J. Baldridge, S. Chatterjee, A. Palmer, and B. Wing. 2007. DotCCG and VisCCG: Wiki and programming paradigms for improved grammar engineering with OpenCCG. In T.H. King and E.M. Bender, editors, *GEAF 2007*, Stanford, CA. CSLI.

T. Baldwin, J. Beavers, E.M. Bender, D. Flickinger, Ara Kim, and S. Oepen. 2005. Beauty and the beast: What running a broad-coverage precision grammar over the BNC taught us about the grammar — and the corpus. In S. Kepser and M. Reis, editors, *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, pages 49–70. Mouton de Gruyter, Berlin.

J.A. Bateman, I. Kruijff-Korbayová, and G.-J. Kruijff. 2005. Multilingual resource sharing across both related and unrelated languages: An implemented, open-source framework for practical natural language generation. *Research on Language and Computation*, 3(2):191–219.

E.M. Bender and D. Flickinger. 2005. Rapid prototyping of scalable grammars: Towards modularity in extensions to a language-independent core. In *IJCNLP-05 (Posters/Demos)*, Jeju Island, Korea.

E.M. Bender, D. Flickinger, and S. Oepen. 2002. The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In J. Carroll, N. Oostdijk, and R. Sutcliffe, editors, *Proceedings of the Workshop on Grammar Engineering and Evaluation, COLING 19*, pages 8–14, Taipei, Taiwan.

E.M. Bender. 2007. Combining research and pedagogy in the development of a crosslinguistic grammar resource. In T.H. King and E.M. Bender, editors, *GEAF 2007*, Stanford, CA. CSLI.

D.G. Bobrow, C. Condoravdi, R.S. Crouch, V. de Paiva, L. Karttunen, T.H. King, R. Nairn, L. Price, and A Zaenen. 2007. Precision-focused textual inference. In *ACL-PASCAL Workshop on Textual Entailment and Paraphrasing*, Prague, Czech Republic.

M. Butt, H. Dyvik, T.H. King, H. Masuichi, and C. Rohrer. 2002. The parallel grammar project. In J. Carroll, N. Oostdijk, and R. Sutcliffe, editors, *Proceedings of the Workshop on Grammar Engineering and Evaluation at COLING 19*, pages 1–7.

A. Copestake, D. Flickinger, C. Pollard, and I.A. Sag. 2005. Minimal recursion semantics: An introduction. *Research on Language & Computation*, 3(2–3):281–332.

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

C. Donohue and I.A. Sag. 1999. Domains in Warlpiri. Paper presented at HPSG 99, University of Edinburgh.

S. Drellishak and E.M. Bender. 2005. A coordination module for a crosslinguistic grammar resource. In Stefan Müller, editor, *HPSG 2005*, pages 108–128, Stanford. CSLI.

D. Flickinger and E.M. Bender. 2003. Compositional semantics in a multilingual grammar resource. In E.M. Bender, D. Flickinger, F. Fouvry, and M. Siegel, editors, *Proceedings of the Workshop on Ideas and Strategies for Multilingual Grammar Development, ESSLLI 2003*, pages 33–42, Vienna, Austria.

D. Flickinger. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering*, 6 (1):15 – 28.

L. Hellan and P. Haugereid. 2003. NorSource: An exercise in Matrix grammar-building design. In E.M. Bender, D. Flickinger, F. Fouvry, and M. Siegel, editors, *Proceedings of the Workshop on Ideas and Strategies for Multilingual Grammar Development, ESSLLI 2003*, pages 41–48, Vienna, Austria.

R. Kim, M. Dalrymple, R.M. Kaplan, T.H. King, H. Masuichi, and T. Ohkuma. 2003. Multilingual grammar development via grammar porting. In E.M. Bender, D. Flickinger, F. Fouvry, and M. Siegel, editors, *Proceedings of the Workshop on Ideas and Strategies for Multilingual Grammar Development, ESSLLI 2003*, pages 49–56, Vienna, Austria.

T.H. King, M. Forst, J. Kuhn, and M. Butt. 2005. The feature space in parallel grammar writing. *Research on Language and Computation*, 3(2):139–163.

A. Kinyon, O. Rambow, T. Scheffler, S.W. Yoon, and A.K. Joshi. 2006. The metagrammar goes multilingual: A cross-linguistic look at the V2-phenomenon. In *TAG+8*, Sydney, Australia.

V. Kordoni and J. Neu. 2005. Deep analysis of Modern Greek. In K-Y Su, J. Tsujii, and J-H Lee, editors, *Lec-ture Notes in Computer Science*, volume 3248, pages 674–683. Springer-Verlag, Berlin.

F. Lareau and L. Wanner. 2007. Towards a generic multilingual dependency grammar for text generation. In T.H. King and E.M. Bender, editors, *GEAF 2007*, pages 203–223, Stanford, CA. CSLI.

J.T. Lønning and S. Oepen. 2006. Re-usable tools for precision machine translation. In *COLING|ACL 2006 Interactive Presentation Sessions*, pages 53 – 56, Sydney, Australia.

M. Marimon, N. Bel, and N. Seghezzi. 2007. Test-suite construction for a Spanish grammar. In T.H. King and E.M. Bender, editors, *GEAF 2007*, Stanford, CA. CSLI.

Stefan Müller. 2007. The Grammix CD-ROM: A software collection for developing typed feature structure grammars. In T.H. King and E.M. Bender, editors, *GEAF 2007*, Stanford, CA. CSLI.

R. Nordlinger. 1998. *A Grammar of Wambaya, Northern Australia*. Research School of Pacific and Asian Studies, The Australian National University, Canberra.

S. Oepen, E.M. Bender, U. Callmeier, D. Flickinger, and M. Siegel. 2002. Parallel distributed grammar engineering for practical applications. In *Proceedings of the Workshop on Grammar Engineering and Evaluation, COLING 19*, Taipei, Taiwan.

S. Oepen, D. Flickinger, K. Toutanova, and C.D. Manning. 2004. LinGO Redwoods. A rich and dynamic treebank for HPSG. *Journal of Research on Language and Computation*, 2(4):575 – 596.

Stephan Oepen, Erik Velldal, Jan Tore Lnning, Paul Meurer, Victoria Rosn, and Dan Flickinger. 2007. Towards hybrid quality-oriented machine translation. On linguistics and probabilities in MT. In *TMI 2007*, Skvde, Sweden.

C. Pollard and I.A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. CSLI, Stanford, CA.

S. Robinson, G. Aumann, and S. Bird. 2007. Managing fieldwork data with Toolbox and the Natural Language Toolkit. *Language Documentation and Conservation*, 1:44–57.

I.A. Sag. 1997. English relative clause constructions. *Journal of Linguistics*, 33(2):431 – 484.

M. Siegel and E.M. Bender. 2002. Efficient deep processing of Japanese. In *Proceedings of the 3rd Workshop on Asian Language Resources and International Standardization, COLING 19*, Taipei, Taiwan.

K. Toutanova, C.D. Manning, D. Flickinger, and S. Oepen. 2005. Stochastic HPSG parse selection using the Redwoods corpus. *Journal of Research on Language and Computation*, 3(1):83 – 105.

E. Velldal. 2007. *Empirical Realization Ranking*. Ph.D. thesis, University of Oslo, Department of Informatics.