Meta-modeling of Tense and Aspect in a Cross-linguistic Grammar Engineering Platform

Laurie Poulson

Abstract

This work describes the implementation of a meta-modeling approach to morphologically marked tense and aspect within the LinGO Grammar Matrix customization system. The Matrix customization system creates a small HPSG grammar for a language based on answers to questions about particular linguistic phenomena. In the original system, these answers indicated a choice of pre-determined analyses of the various phenomena. The complexity of tense and aspect systems, the cross-linguistic variation in those systems, as well as the variation in linguists' analyses of tense and aspect and their underlying assumptions, make tense and aspect difficult to implement in this way. Meta-modeling is proposed as an answer to the question of how tense and aspect can be implemented in the customization system while imposing a minimal amount of analysis on the linguist-user. The meta-modeling approach supports the linguist-user in developing their own analytical model of tense and aspect for a particular language. Within the framework of HPSG, as implemented in the LinGO Grammar Matrix, the meta-modeling of tense and aspect provides flexibility for the linguist-user with regards to the definition of types, hierarchies and features. In addition, this approach allows the user to determine how tense and aspect features relate to specific morphemes as well as how the features contribute to the definition of grammatical constraints.

Contents

1 Introduction

The Grammar Matrix project (Bender et al., 2002; Bender & Flickinger, 2005) is a cross-linguistic grammar engineering project designed to support the development of computer-implemented grammar models. These grammars are not descriptive grammars in the sense of Rice (1989) but instead computer models of grammars like the English Resource Grammar (ERG; Flickinger (2000)) that can parse input sentences and generate from semantic representations. Matrix style grammars consist of several files that work together to define the grammar model. These grammars are compatible with the DELPH-IN suite of software tools, including the LKB grammar development environment (Copestake, 2001).

The Matrix customization system produces small but working grammars customized to a linguist's specifications with regards to particular language phenomena. These grammars are designed to be further developed by grammar engineers. Phenomena are added to the customization system incrementally. This work describes the initial implementation of the phenomena of tense and aspect in the customization system. This implementation specifically tackles the processing of morphologically[1] marked verbal tense and aspect elements. The intention of this implementation is not to provide a temporal interpretation of sentences but rather to model syntactic constraints on grammaticality and to produce semantic representations usable as input for temporal interpretation.

The Grammar Matrix customization system produces small grammars reflecting 'choices', i.e., answers to questions elicited through a typologically-informed questionnaire. Prior to the development of this tense and aspect implementation, the basic approach had been to develop a set of analyses for each phenomenon. Choices on the questionnaire determined which of the analyses were included in the grammar for that language. This approach motivates the use of the term *library* to refer to the implementations of phenomena in the customization system. For example, the coordination library (Drellishak & Bender, 2005), based on typological research, implements pre-defined analyses for an impressive range of strategies languages employ for and-coordination[2] The only user-defined piece of the customized grammar that results from the questionnaire choices is the spelling of the coordination marker.

I am not arguing here that the coordination library is deficient but instead that there are characteristics of particular phenomena that make them amenable to this approach. Phenomena that are fairly well understood, for which there is general agreement about the elements involved and the role those elements play, and that have cross-linguistically consistent semantics, lend themselves to an approach utilizing a library of pre-defined analyses. In contrast, there are characteristics of tense and aspect that make it difficult to model in terms of a collection of developer-defined analyses. The characteristics are discussed in some detail in §3. For now, let me simply state the research question: How can the customization system be extended to handle tense and aspect marking without imposing a one-size fits all

[1]The term *morphological* is often used to refer exclusively inflectional elements. This meaning is used in the context of a contrast between synthetic (morphological) vs. analytic forms. One the other hand, *morphological* may also be used more generally to refer to any morpheme, including independent lexemes. Throughout this work, I use the term with the latter, more general, meaning.

[2]The coordination library covers and-type coordination of the same (or similar) grammatical categories.

analysis of tense and aspect on the linguist-user? In answer to that question, I assume a meta-modeling approach that allows the linguist-user quite a bit of flexibility to define the basic elements of the tense and aspect system for the language they are describing. Specifically, instead of making choices that lead to a pre-defined analysis, users define the basic tense and aspect elements according to their own analysis.

In the following section (§2), I provide a bit more detail about the Grammar Matrix. In §3, I discuss tense and aspect, particularly in terms of variations among languages and differences in analyses. In §4, I discuss more fully the motivation for a meta-modeling approach as well as the overall scope of the current implementation. The details of the implementation are provided in §5, including an extended example that utilizes multiple components of the implementation, and a discussion of future work. Finally, I conclude in §6 with some brief remarks about the implementation.

2 The Grammar Matrix

The Grammar Matrix project is fundamentally a grammar engineering project; phenomena are identified and implemented through an incremental process of development and testing. The grammar engineering workflow, as schematized in Bender et al. (2011), is illustrated in Figure 1. The process of creating a grammar for a language begins with data which, in Bender's schematic, is in the form of an initial test suite. A phenomenon is identified in the data, an analysis of the phenomenon is implemented, tested and refined, and then the process is repeated for the next phenomenon. Each new analysis must be compatible with previous elements in the implementation. This may lead to extensive debugging and even re-analysis of the current or previous phenomena.

Grammar engineering is a very time-intensive process. The Grammar Matrix project was conceived to jump-start grammar creation by providing a foundation on which to build grammars. Grammar Matrix style grammars are written in TDL (Type Description Language) (Krieger & Schäfer, 1994) within the framework of Head-driven Phrase Structure Grammar (HPSG) (Pollard & Sag, 1994; Sag et al., 2003) incorporating Minimal Recursion Semantics (MRS) (Copestake et al., 2005). The original Grammar Matrix, referred to in the context of the customization system as the "Matrix core grammar", contains those elements believed to be cross-linguistically applicable for HPSG grammars with MRS style semantic representations. This core contains such elements as general types associated with the basic feature geometry, types associated with the semantics, types for basic constructions such as head-complement, head-specifier etc., and general classes of rules including those that implement the general principles of HPSG (Bender et al., 2002).

The development of the customization system for the Grammar Matrix began with the observation that the same analysis of a phenomenon is often appropriate for multiple languages. For example, many languages share the same basic word order, several languages form questions through subject-verb inversion and, in many languages, sentential negation is a verbal suffix. The customization system initially leveraged previous analyses of common variations of word order, negation and yes/no questions (Bender & Flickinger, 2005). Since

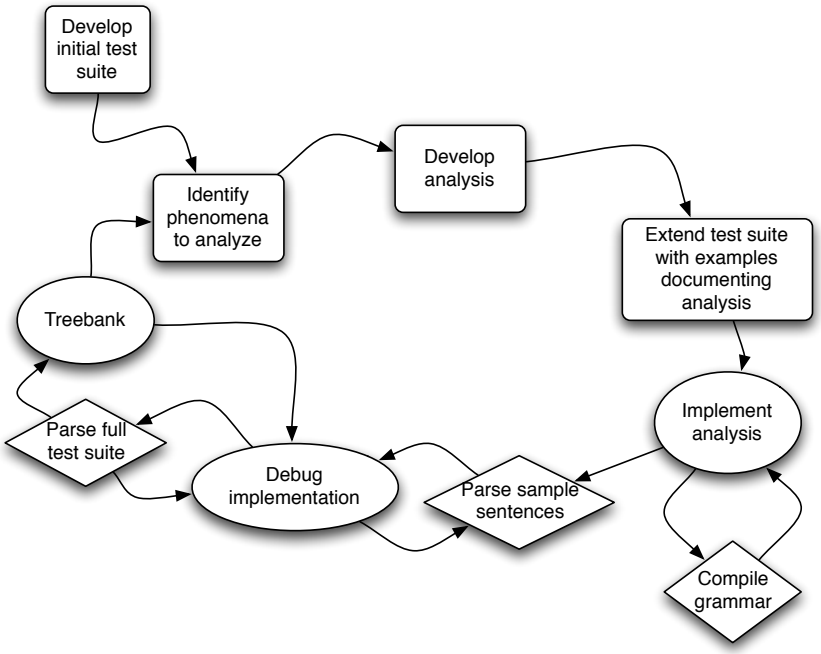


Figure 1: Grammar engineering

then, additional phenomena have been added based on typological research: coordination (Drellishak & Bender, 2005), case and agreement (Drellishak, 2009) and tense and aspect as described in this paper. In addition, projects have been undertaken to upgrade the handling of word order (Fokkens, 2010) and argument drop (Saleem & Bender, 2010).

The schematic diagram in Figure 2, reproduced from Bender et al. (2010), illustrates the Grammar Matrix customization system while abstracting away from some of the details of the actual customizing step. The user provides answers to questions on the questionnaire. The questionnaire is checked for consistency and the choices made by the user are extracted. The 'stored analyses' in Figure 2 abstractly represent the phenomena-specific information spread throughout the customization code. The answers from the questionnaire and the analyses from these 'libraries' inform the language-specific portions of the grammar. This language-specific information, combined with the Matrix core, constitutes the starter-grammar. This small starter-grammar can both parse and generate.[3] The starter-grammar is customized for the linguist-user's language, based on the information provided on the questionnaire. For linguist-users interested in building broad-coverage implemented grammars, this can represent a significant step. In addition, a starter-grammar may be useful for linguists interested in testing linguistic analyses within the context of a small implemented grammar.

Figure 3 contains a screenshot of most of the main page of the questionnaire. Each section of the questionnaire contains a variety of questions related to particular phenomena. Additional details about the questionnaire and the customization system as it relates to tense and aspect are described in §5. However, first, I provide some background on the

[3]The Matrix customization project uses the LKB (Linguistic Knowledge Builder) (Copestake, 2001) grammar development environment for parsing and generating as well as for debugging.

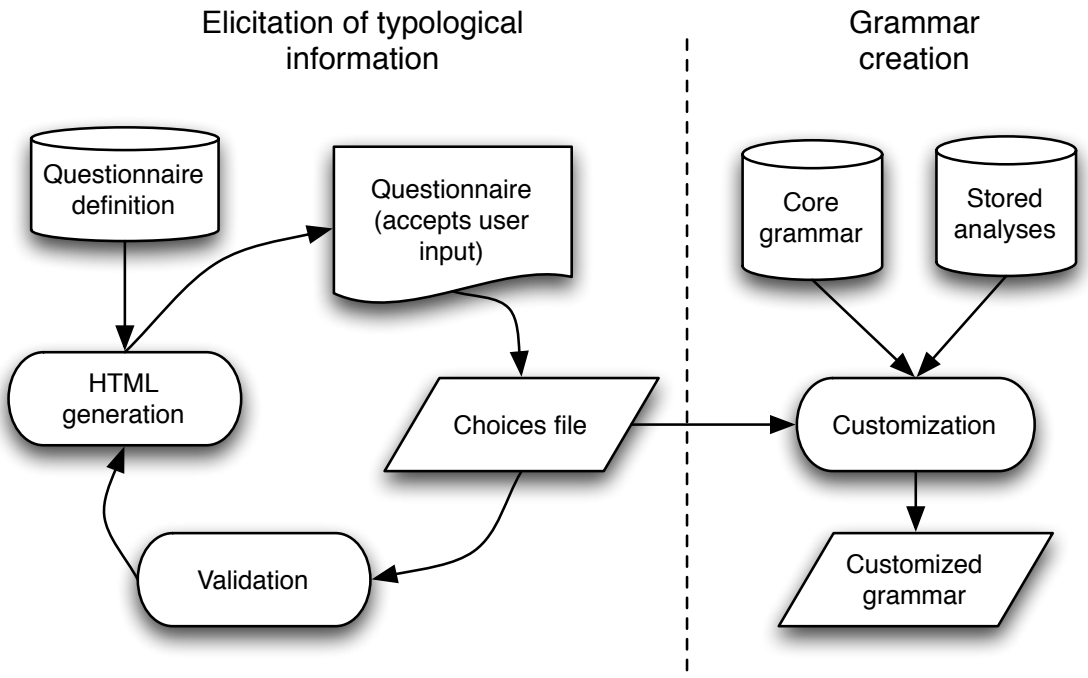


Figure 2: The Grammar Matrix customization system

phenomena of tense and aspect.

3 Tense and Aspect

> In any utterance a peculiar importance is universally attached to the temporal contour of a state or action and the speaker's attitude towards it. The grammatical correlates of these contours and attitudes are the categories of Tense, Aspect, and Modality; they are pervasive, they are universal (in that no language lacks all three), and every speech event must incorporate one or more of them. (Hopper, 1982)

Given the universality of tense, aspect and mood/modality[4] (TAM), there is strong interest in providing support for TAM in the Matrix. Givón (1984), however, refers to TAM as "one of the most complex subsystems of the grammar". TAM-systems function across the interfaces of syntax, semantics and morphology as well as pragmatics and discourse. This implementation does not address the full tense and aspect system but begins the process of supporting the modeling of TAM with a preliminary implementation of marked tense and aspect.[5]

[4]The distinction between mood and modality is beyond the scope of this work.

[5]Extending the implementation to mood/modality is future work. In addition, discussion of TAM-systems should probably also include evidentials—whether they are (see Palmer, 1986) or are not (see de Haan, 1999) subsumed by modality—but evidentials also lie outside of the scope of this work.

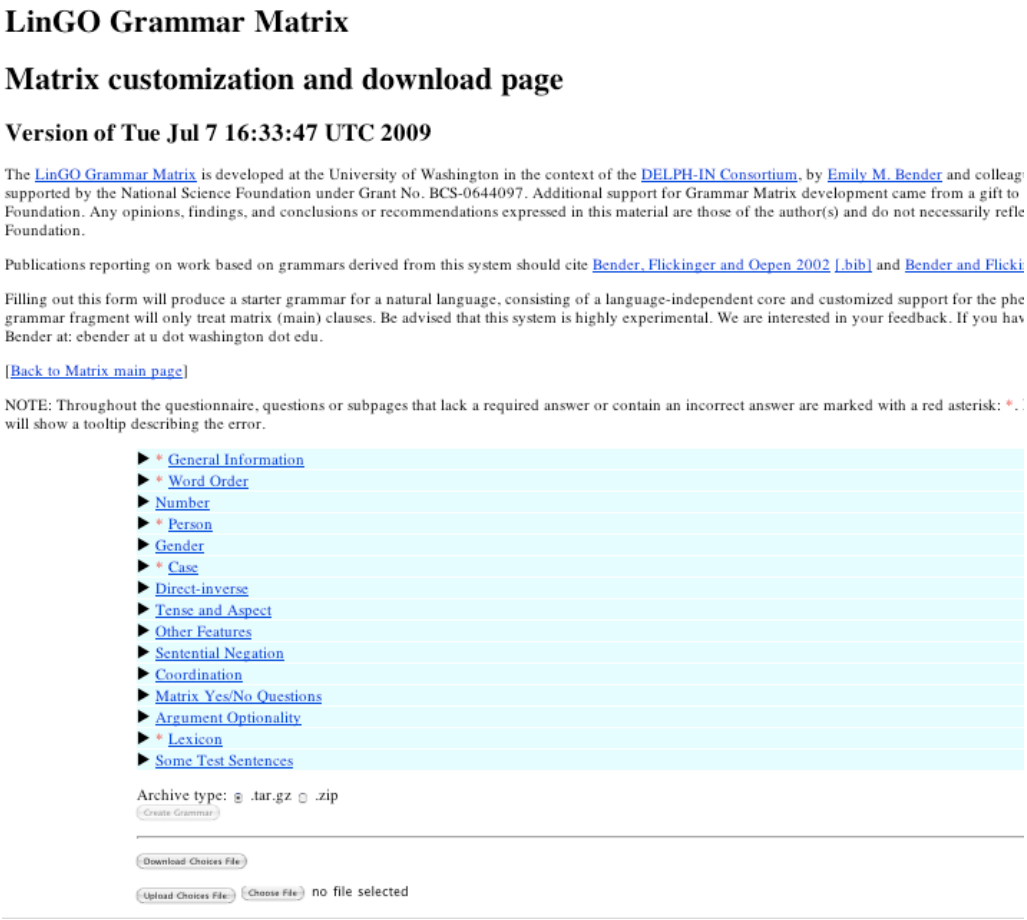


Figure 3: Questionnaire: main page

There is a vast literature on tense and aspect as is evidenced by the size of the bibliography Binnick (2006) has amassed related to tense and aspect.[6] In this paper I will not attempt to provide a full review of this subject. However, in §3.1 and §3.2, I discuss some characteristics of tense (§3.1) and aspect (§3.2) and how they vary cross-linguistically.

3.1 Tense

> The grammatical forms known as tenses establish the time talked about in a sentence, in conjunction with other temporal expressions. ... Temporal interpretation is determined by interpretive rules that take as input the surface structure of a sentence and the context in which it appears. (Smith, 2007)

Tense is but one of many factors that contribute to the temporal interpretation of a sentence. Comrie (1985, p.9) defines tense as "grammaticalised expression of location in

[6]Currently, the bibliography contains around 9000 entries.

time". Tense may locate situations[7] as simultaneous with the deictic center or at some distance before or after it, where the deictic center may be the present moment (absolute tense) or some other point in time (relative tense) (Comrie, 1985, p.1,9-10).

Most descriptions of the semantics of tense assume three important points. These points are generally indicated by some version of S (speech time), E (event time) and R (reference time), based on Reichenbach (1947).[8] A three point system is useful for describing absolute-relative tense (Comrie, 1985). Absolute-relative tenses are described by Comrie (ibid., p.65) as those "determined by a reference point being before or after the present moment, and by the situation being located before or after that reference point". The English Pluperfect,[9] e.g., *He had eaten* is an example of an absolute-relative tense. This implementation addresses absolute tense so the following discussion focuses on that (see §4.2 for more on the scope of this implementation).

3.1.1 Dimensions of variation

Comrie (ibid.) defines two dimensions of variation for absolute tense: direction and distance from the present moment (speech time). The English [iso: eng][10] examples in (1)[11] illustrate what is meant by direction. Example (1a), with present tense marked on the finite verbal element, describes an event simultaneous with the speech time while example (1b) describes an event prior to the speech time. In contrast, a future tense would mark an event subsequent to the speech time.

(1) a. He is walking.
He AUX.PRS walk-PROG [eng]

b. He was walking.
He AUX.PST walk-PROG [eng]

The distance dimension can be illustrated by the Cocama (Cocama-Cocamilla, iso: cod) examples in (2) from Faust (1978, p.42), cited in Bybee et al. (1994, p.98).[12] All of these examples describe events prior to speech time but the temporal distance from the event

[7]The term *situation*, coined by Comrie (1976) is a commonly used cover term for both states and events leaving the term *event* for referring to non-stative situations.

[8]Note, however, that Klein (1994) argues for a slightly different three point system.

[9]In this work, I use the notational convention from Comrie (1976): When tense (or aspect) terms refer to language specific categories or forms, they are capitalized, e.g., English Present Progressive. On the other hand, when terms (often the same terms) are used to indicate a semantic notion or a cross-linguistic category, they are lower case.

[10]The ISO 639-3 language codes (iso) used in this document are from Lewis (2009).

[11]The gloss lines appear as they do on these examples for clarity regarding the tense direction. However, they are somewhat misleading regarding the progressive aspect since it is the auxiliary verb construction (aux+*walk*) that defines the progressive, not any marking on the verb *walk*.

[12]Examples in this paper that have been borrowed from other works may have slight alterations as I may have changed the abbreviations in the morpheme gloss lines to be consistent across the paper. Appendix A defines the gloss abbreviations used.

to the speech time varies. Three different past time frames are described in these examples: Hodiernal (related to today), Hesternal[13] (related to yesterday), and Remote.[14] While the English translations of these sentences utilize adverbials to make the distinctions, in Cocama, the distinctions are marked by obligatory tense inflection (ibid.). According to Dahl & Velupillai (2008b), as many as one fifth of the 100 languages in their sample exhibit remoteness distinctions in past or future tenses.

(2) a. Ritama-ca tuts-ui.
town-to go-PST(Hodiernal)
'I went to town today.' [cod]

b. Ritama-ca tutsu-icuá.
town-to go-PST(Hesternal)
'I went to town yesterday.' [cod]

c. Ritama-ca tutsu-tsuri.
town-to go-PST(Remote)
'I went to town a long time ago.' [cod]

Tense is a grammatical category and as such is distinct from lexical temporal expressions such as temporal adverbials. Tense, for example, is generally obligatory even when the information conveyed is redundant. For example, in the sentence *She walked yesterday.*, although the adverb *yesterday* clearly provides the information that the event occurred in the past, the tense is still required.

However, Comrie (1985, p.10) notes that the line between what is grammatical and what is lexical is often indistinct. For example, the English expression *to be about to V* places a situation in the future but is not clearly grammaticalized (Comrie, 1985, p.95). Since grammaticalization is a diachronic process, there is a natural cline from lexical to grammatical. Bybee et al. (1994) argues that various degrees of grammaticalization and residue from original lexical meanings account for the multiple uses, nuanced meanings and restricted distribution of individual tense and aspect morphemes. Maltese, for example has a future form *se* (*ser, sa, sejjer* 'going') that, according to Ebert (2000), is often characterized as an imminent or definite future and does not combine with stative verbs. Ebert suggests that the original intentional meaning of the morpheme accounts for its ungrammaticality with statives.

There is one more thing to note regarding the examples in (2): The meaning of language specific categories, such as Hodiernal or Remote, though found in many languages, are in fact language specific. The meaning associated with the category depends on the tense system context. For example, while Remote refers to *before yesterday*, in Cocama, it may refer to *before this year* or *before this generation* in another language, depending on the set of tense distinctions in that language. Universal, i.e., cross-linguistically valid, tense (and aspect) categories have been identified (Dahl, 1985; Bybee et al., 1994) however, these are based

[13]The term *hesternal* was coined by Dahl (1985).

[14]Note that I have added the distinguishing parenthetical labels to these examples.

on prototypes. Particular tense forms, and the meanings these convey, are language specific (Dahl, 2000). The language-specific nature of categories is addressed again in §3.2.1.

3.1.2 Grammaticalization

Of the possible tense distinctions related to distance or direction, languages differ with respect to which are grammaticalized. According to Dahl & Velupillai (2008a), as well as Lindstedt (2001), Finnish (iso: fin) has no future tense.[15] In Finnish, as seen in example (3),[16] both the present (today) and the future (tomorrow) utilize the same verb form. This is in contrast to the French (iso: fra) equivalents in example (4) where the verbs have distinct present and future forms.

(3) a. Tänään on kylmää.
today is cold-PART
'It is cold today.' [fin]

b. Huomenna on kylmää.
tomorrow is cold-PART
'It will be cold tomorrow.' [fin]

(4) a. Il fait froid aujourdhui.
it do.PRS.3SG cold today
'It is cold today.' [fra]

b. Il fera froid demain.
it do.FUT.3SG cold tomorrow
'It will be cold tomorrow.' [fra]

French is an example of a language with a three way primary tense contrast: past/present/future while the basic contrast in Finnish is past/non-past. English is also classified by many, if somewhat controversially, as a past/non-past language. The controversy often centers on the question of whether or not the future meaning of *will* is subsumed by modality.[17] In contrast, the Hopi (iso: hop) tense system is based on a non-future/future distinction (with various degrees of remoteness) (Frawley, 1992, p.339).[18]

Furthermore, there are languages with little or no tense at all. Mandarin (iso: cmn) and Thai (iso: tha) (Smith, 2005), as well as Burmese (iso: mya) and Dyirbal (iso: dbl) (Comrie, 1985), are examples of languages with no tense. Smith (2005) argues that, in such languages, "aspect allows inference about temporal location". For example, in Mandarin, bounded situations are interpreted as past events by default while ongoing situations are interpreted

[15] Although see Thieroff (2000) for a different opinion regarding whether Finnish has a future tense.

[16] The examples in (3) and (4) are from Dahl & Velupillai (2008a).

[17] See this posting on Language Log by Geoffrey Pullum: `http://itre.cis.upenn.edu/~myl/languagelog/archives/005471.html` and this interchange on LinguistList: `http://www.linguistlist.org/issues/8/8-178.html` for some detail of the issue and an interesting discussion.

[18] Frawley credits Malotki (1983) with the analysis of the Hopi tense system.

in the present (Smith, 2005). Even in languages with tense, aspect can play a larger role in locating situations in time than tense (Smith, 1997, 2005). In Russian, bounded situations (perfectives) are interpreted in the future unless specifically marked as past (Comrie, 1976, p.66-67).[19] This future reading is inferred in the context of a past/nonpast tense contrast since bounded situations cannot be contained in the present. However, it can be difficult to differentiate semantic content contributed grammatically from that which is pragmatically inferred.

3.1.3 Form and meaning

In addition to varying with respect to what is grammaticalized, tense marking has various morphological forms. Tense is, primarily, a verbal category[20] and, cross-linguistically, is frequently marked by verbal inflection. For example, English tense inflection appears on finite verb forms. The simple Past tense is marked on the main verb as in example (5a) while in Progressives, it is marked by inflection on the suppletive auxiliary as in (5b).[21]

(5) a. He walked.
he walk-PST [eng]

b. He was walking.
he AUX.PST walk-PROG [eng]

However, periphrastic expressions, especially auxiliary verb constructions, are also quite common for marking tense. Comrie (1985, p.11), drawing on data from Hyman (1980), cites the Niger Congo language Bamileke-Dschang (a.k.a. Yembe, iso: ybb) as an example of a language that marks tense with auxiliaries. The auxiliaries listed in example (6) are obligatory tense markers that do not function as independent lexical items. Many languages have both inflectional and periphrastic tense markers, e.g., English (inflectional past, periphrastic future) and French (both inflectional and periphrastic past, inflectional future).

(6) a. à kè táŋˈˀŋ
he AUX(yesterday) bargain
'He bargained yesterday.' [ybb]

b. à lè táŋˈˀŋ
he AUX(some days ago) bargain
'He bargained some days ago.' [ybb]

Tense (and aspect) markers have both a meaning component and a form component although there is not necessarily a one-to-one correspondence between the two. While markers

[19] For imperfectives there is a periphrastic future tense available.

[20] There are phenomena that have been analyzed as nominal tense (see Nordlinger & Sadler, 2004). These phenomena are beyond the scope of this current implementation.

[21] In the gloss, I have ignored the number agreement inherent in the auxiliary form as it is not germane to the specific point at hand. Also, the PROG (progressive) gloss marking in this example is somewhat misleading in that *-ing* on the verb alone does not mark Progressive; it is marked by the complex construction, aux+Ving.

may convey only a single semantic notion, they may also be in many-to-one, many-to-many or one-to-many relationships between the morphemes of the marker and the content of the semantics. For example, French Imparfait, as illustrated in example (7), is often referred to as a 'tense'. However, the *ait* inflection on the verb conveys not only past tense but also imperfective aspect and 3rd person singular agreement with the subject. Fusional morphemes like this one represent one-to-many morpheme-to-content relationships. In fact, combinations of tense, aspect, modality, evidentiality and agreement are commonly conveyed by single morphemes.

(7) Elle march-ait.
FEM.3SG walk-IPFV.PST.3SG
'She walked / was walking / used to walk.' [fra]

The auxiliary verb construction, Passé Composé, in example (8), on the other hand, consists of two morphemes. In spoken French, Passé Composé may convey either a past perfective or a perfect meaning.[22] In addition, the auxiliary construction has two morphemes but it is not the case that each morpheme independently contributes a single feature or unit of meaning. The form of the auxiliary in Passé Composé is Present Tense; it is precisely in composition with the participial form of the verb that either a past or an absolute-relative past (Present Perfect), is conveyed. The French Passé Composé illustrates a many-to-one relation. In fact, auxiliary verb constructions in general, tend to form many-to-one form-to-content relations.[23]

(8) Elle a march-é.
FEM.3SG AUX.3SG walk-PFV.PST
'She walked.' [fra]

Imparfait and Passé Composé are contrasting past constructions in spoken French. However, while they are routinely referred to as *tenses*, the contrast is actually an aspectual one.[24] This illustrates a source of potential confusion since the term *tense* actually refers to a collection of interrelated items: the grammatical category, the language specific markers and the semantic notions conveyed by the markers. In addition, the term *tense* is often used to describe a morpheme slot, i.e., an affixal position. In that case all morphemes sharing a particular morphological slot, including those contributing, for example, only aspect or modality may be referred to as *tense* morphemes.[25] Much of this may be residual traditional

[22]The Passé Composé is used to convey either a past perfective (simple past) or a perfect. I assume that the perfect constitutes a separate grammatical construction with distinct characteristics. Details of the perfect are beyond the scope of this work. The details presented here pertain to the Passé Composé as simple past.

[23]This can lead to the glossing confusion in example (8); the past tense and perfective aspect is conveyed by the complex construction, not strictly by marking on the main verb. The gloss implies otherwise. This is essentially the same gloss issue footnoted with reference to example (5b).

[24]See §3.2 for more on this distinction.

[25]The same confusion exists for the term *aspect*. For example, Hoxie (1996) describes the "aspects" of Cherokee (iso: chr) as present, imperfective, perfective, imperative and infinitive. From the description, it seems that, at least, present, imperative and infinitive are likely not marking semantic aspect.

usage or for convenience although it may also reflect the issue already mentioned that it is often difficult to determine what is semantically contributed and what is pragmatically inferred.

3.1.4 Interaction

Tense and aspect are integrally linked. They both contribute to the temporal characterization of a sentence. For example, what one language conveys with past tense another might convey with perfective aspect. Often past tense is pragmatically inferred given perfective aspect, not as an entailment but as an implicature. The involvement of pragmatics can complicate these analyses. Dryer (2008) notes that "different descriptions of the same language often differ in whether they characterize a category as one of tense or as one of aspect."

Tense also interacts with other grammatical systems. Case, for example, often interacts with tense. For example, split-case systems may be conditioned on tense (or aspect)[26] values. Comrie (1978) cites data from Georgian (iso: kat), reproduced here as example (9), illustrating conditioning on tense. A nominative/accusative system is required with the Present tense form in example (9a) while the Aorist (past tense) form in example (9b) requires an ergative/absolutive system.

(9) a. Sṭudenṭ-i ċeril-s ċer-s.
student-NOM letter-ACC write-PRS
'The student writes the letter.' [kat]

b. Sṭudenṭ-ma ċeril-i daċer-a.
student-ERG letter-ABS write-PST
'The student wrote the letter.' [kat]

In summary, tense is a grammatical category which interacts extensively with aspect and various other grammatical phenomena. It has two basic dimensions of variation: direction and distance. Within these dimensions, languages vary with respect to what is grammaticalized. The form tense marking takes also varies within and across languages. The most common forms of tense marking are verbal inflection and auxiliary verb constructions. How semantic features relate to these forms can be difficult to ascertain as there is often not a one-to-one relation between form and meaning. In addition, it can be difficult to determine what semantic content is grammaticalized as tense and what is pragmatically inferred.

3.2 Aspect

> It has become commonplace to introduce works on aspect with the remark that there is hardly another field in linguistics so much plagued by terminological and notional confusion. The semantics of time has served as a playground for mental exercise to many generations of philologists, linguists, philosophers, and logicians,

[26] Sherpa (iso: sxr) provides an example of case conditioned on aspect. According to Givón (1984), Sherpa is a split-ergative language exhibiting ergative/absolutive morphology only in the perfective.

> resulting in an impenetrable thicket of definitions, theories, and models. (Sasse, 2002)

According to Smith (1997, p.xiii), "aspectual meaning contributes temporal information and point of view to sentences. It is through aspect that we grasp the type of situation talked about, from a temporal perspective which focuses all or part of the situation. Temporality in this sense concerns the way situations unfold in time." According to Sasse (2002), aspect is fundamentally about situations being bounded or unbounded. For example, the French examples (7) and (8), repeated here as (10), illustrate a contrast in the speaker's description of a past event. Example (10a) describes a walking event including the boundaries (endpoints), i.e., as a completed unit, while example (10b) describes a walking event process ignoring endpoints. This contrast is reflected in a grammatical distinction imposed by the speaker that controls the focus or viewpoint on the situation.

The English examples in (11) illustrate another type of bound. Example (11a) describes an inherently bound event, i.e., an event with an inherent endpoint that may be reached. The building event has an inherent endpoint, i.e., the building ends when the boat is complete. In contrast, the situation in example (11b) has no inherent endpoint but instead reflects an ongoing state. This bound/unbound contrast results from inherent temporal properties of the situations described.

(10) a. Elle a march-é.
FEM.3SG AUX.3SG walk-PFV.PST
'She walked.' [fra]

b. Elle marchait.
FEM.3SG walk-IPFV.PST.3SG
'She walked / was walking / used to walk.' [fra]

(11) a. She built a boat.
she build-PST a boat [eng]

b. She loves it.
she love-PRS it [eng]

These distinct ways of bounding situations, grammatical and inherent, comprise what Sasse (ibid.) describes as two dimensions of aspect. The aspectual content of a sentence depends on the interaction between these two independent components or dimensions. Some commonly used terminology for describing basically these two dimensions includes grammatical vs. lexical, outer vs. inner or aspect vs. Aktionsart, although most of these terms have quite a wide range of meaning. Sasse calls the two dimensions ASPECT$_1$ and ASPECT$_2$ respectively however, in this work, I use the terminology from Smith (1991, 1997), referring to them as viewpoint and situation aspect. Viewpoint aspect refers to the grammatical contrast that Comrie (1976, p.3) describes as "different ways of viewing the internal temporal constituency of a situation". The primary viewpoint contrast is between perfective and imperfective. Situation aspect refers to inherent temporal characteristics of situations, such as stativity, durativity and telicity.

Cross-linguistically, languages vary widely with respect to aspect. Within the bi-dimensional aspect model, languages may vary in terms of the emphasis placed on one or the other dimension. For example, Slavic languages have strong viewpoint systems while Germanic languages little or no marked viewpoint aspect and an elaborated situation aspect system. In addition, which aspectual notions are grammatically distinguished, how they are morphologically or syntactically expressed and how the expression of aspect interacts with other grammatical systems, can be quite language-specific. In §3.2.1 and §3.2.2, I discuss aspect in terms of the two dimensions, viewpoint and situation, and provide some examples of cross-linguistic variation.

3.2.1 Viewpoint aspect

Viewpoint aspect conveys the speaker's subjective perspective on a situation, describing the situation from a particular viewpoint or focusing on particular aspects or elements of the situation. Smith (1991, 1997) divides situations into endpoints and stages and defines the primary viewpoints as perfective, a view including both endpoints, and imperfective, a view including no endpoints.[27]

Viewpoint aspect, like tense, is commonly marked through verbal inflection but may also be marked by auxiliaries, either with, or without, specific complement forms. French, as illustrated in example (10), utilizes both inflection and auxiliary constructions to mark viewpoint. Viewpoint, unlike tense, may also be marked on lexical items through derivational rather than inflectional processes.

For example, in Russian, verbs fall into one of two related verb classes: Perfective or Imperfective. Perfective verbs are derivationally obtained from Imperfective stems (Binnick, 1991). In Russian, aspect is obligatorily marked. The use of a Russian Perfective verb conveys a situation as a single completed action. This interpretation is generally analyzed as resulting from the contribution of perfective aspect (Sasse, 2002).[28] Example (12) illustrates the Russian Perfective/Imperfective distinction.[29] The Imperfective in example (12a) is ambiguous about the completion of the book reading. More generally, a Russian verb in the Imperfective class is ambiguous about viewpoint (Paducheva & Pentus, 2008). In contrast, the Perfective, as in example (12b), unambiguously describes a single complete event; the reader is no longer reading and the book was completely read.

(12) a. Ja ĉita-l knigu.
I read-PST(IPFV) book
'I was reading a book / used to read a book / read a book.' [rus]

b. Ja proĉita-l knigu.
I read-PST(PFV) book
'I read a book.' (an entire book) [rus]

[27]Smith (1991, 1997) also argues for a third primary viewpoint available in some languages: neutral, a view including the initial endpoint and at least one stage.

[28]However, see §3.2.3 for more on this point.

[29]These examples and interpretations are from Anya Dormer (p.c.). Any errors are mine.

Imperfective and perfective are the most commonly marked viewpoint categories. Dahl & Velupillai (2008c) reports that, in a survey of 222 languages, 101 were categorized as having a grammaticalized perfective/imperfective contrast. However, many languages have additional grammaticalized aspects that relate to viewpoint. For example, according to the analysis in Muansuwan (2002), Thai, in addition to Imperfective and Perfective aspect, has the categories Post-Inchoative and Semi-Perfective. Semi-Perfective is described by Koenig & Muansuwan (2000) as similar to the Perfective but without entailing completion of telic events.

Many languages, including English, have a grammaticalized progressive, often analyzed as a subtype of imperfective. French, however, does not. The meaning associated with English Progressive is but one of several interpretations of French Present tense forms, as illustrated in (13a). The Imparfait example in (13b) and the Passé Composé example in (13c) illustrate again the perfective versus imperfective distinction in French. While the French Imparfait can convey a meaning equivalent to an English Progressive, it is, in fact, a broader category including some meanings conveyed in English by the simple Present or Past (e.g., habitual).

(13) a. Il parl-e.
he talk-PRS
'He talks.' / 'He is talking.' [fra]

b. Il parl-ait.
he talk-IPFV.PST
'He was talking.' / 'He talked.' (habitually) [fra]

c. Il a parl-é.
he AUX talk-PFV.PST
'He talked.' (on a specific occasion) [fra]

In addition to variation in grammaticalized categories, the relationship between aspect categories may also vary across languages. For example, perfective and imperfective categories are contrastive in many languages, e.g., Russian and French, and as such are mutually exclusive. However, in Bulgarian there is a form referred to as the Imperfective Aorist used "to indicate an action which is presented as a single whole (whence the Aorist as a marker of perfectivity), but with internal complexity (whence the Imperfective as a marker of imperfectivity)" (Comrie, 1976, p.23).

Viewpoints may also vary in more subtle ways, complicating the notion of a cross-linguistic category. Various analyses of the progressive demonstrate the complication. In general terms, Comrie (1976) defines progressive as a subtype of imperfective expressing an ongoing activity that is only applicable to non-statives. In contrast, Dahl (1985) argues that progressives are distinct from imperfectives, although the categories frequently overlap, and that they may, or may not, be acceptable with statives, depending on the language. Sasse (2002, p.210) concludes that progressive may, or may not, be analyzed as a more specific type of imperfective aspect, depending on the language and the model of aspect assumed. In

language-specific terms, the Thai Progressive is allowed with certain temporary states but with no permanent states (Muansuwan, 2002). The Navajo Progressive is compatible with a very limited number of verbs (Midgette, 1987). Haspelmath (2008), concludes that the meaning associated with the English Progressive construction is specific to English and the precise range of its uses is probably not shared with any other language.

Japanese has a construction, *-te iru*, that also illustrates the language-specific nature of aspectual categories. The *-te iru* construction expresses ongoing situations and in many contexts appears equivalent to the English Progressive. However, Ogihara (1998) describes a significant distinction between *-te iru* and the English Progressive. The difference reveals itself in the context of achievement events as illustrated in (14) taken from Ogihara (1998).[30] The English sentence in example (14a) describes an ongoing situation occurring before John falls asleep. In contrast, (14b) describes an ongoing situation occurring after Taro falls asleep. Thus, although the Japanese *-te iru* construction can convey a progressive meaning, it can also convey a type of resultative meaning not conveyed by the English Progressive.

(14) a. John was falling asleep.
John AUX.PST fall-PROG asleep.
John was approaching sleep (not yet asleep). [eng]

b. Taroo-wa nemuri-ni tui-te i-ta.
Taro-ACC sleep-DAT begin/arrive-*te iru*-PST
'Taro was asleep (as a result of having fallen asleep).' [jpn]

From a typological perspective, various cross-linguistic aspect categories have been proposed (see Dahl, 1985; Bybee et al., 1994). However, as with tense categories, these typological categories refer to prototypical properties. Language-specific categories that share enough of the prototypical properties are assumed to be in the same typological category. Language-specific categories, however, can vary widely not only in how they are marked but also in what they convey. As a final example of this variation: The Russian Perfective always conveys completion. According to Bybee et al. (1994, p.89), this is typical of derivationally marked perfectives. Inflectional perfectives, on the other hand, are not usually completive (ibid.).

Haspelmath (2008) argues that universal comparative concepts are more useful for cross-linguistic comparison than prototype categories; imperfective and perfective, for example, may express universal semantic notions but not language-specific categories. Per Haspelmath (ibid., p.2), "Each language has its own categories, and to describe a language, a linguist must create a set of descriptive categories for it. These categories are often similar across languages, but the similarities and differences between languages cannot be captured by equating categories across languages." Similarly, Johanson (2000, p.45) comments that "languages obviously delimit and divide [the conceptual space of aspectotemporality] differently" and that "grammatical meaning is language-specific".

In summary, viewpoint aspect is a grammatical category that specifies the speaker's focus or viewpoint on a situation. It is often marked through inflection but may also be marked

[30]Note that I added the English gloss in example (14a).

by auxiliary verb constructions or derivational morphology. As with tense, which viewpoint categories are grammaticalized and how they are marked is language-specific. Also language-specific are the relationships between grammaticalized categories and the semantic content the categories convey.

3.2.2 Situation aspect

In contrast to viewpoint aspect which establishes a viewpoint on a situation, situation aspect refers to its temporal properties or characteristics. In other words, these characteristics do not describe the perspective on a situation but rather describe the type of situation it is. The term *situation aspect* overlaps with, or subsumes, lexical aspect, inherent aspect and some uses of the term Aktionsart.

Situation aspect is generally discussed in terms of two separate but interconnected concepts: situation types and situation aspect features. Situation types are prototype classes that organize situations according to their temporal properties. Various classifications of situation types have been proposed, the most common in the literature being those articulated by Vendler (1957): States, Activities, Achievements and Accomplishments. In many models of situation aspect, e.g., that of Smith (1991, 1997) these types are names for collections of temporal features.

Various sets of features have been proposed as the fundamental properties of situation types. Table 1 lists situation types and their defining features according to Smith (1991) as well as English examples of each situation type. States and events are differentiated by whether they are static or dynamic. All events are dynamic. Events break down into Vendler's basic types, Activities, Accomplishments and Achievements, and an additional situation type: Semelfactive (Smith, 1991). Events are differentiated by the features durative vs. instantaneous (punctual), and telic vs. atelic. [31]

States: static	*She is happy.*		
Events: dynamic			
	Accomplishments:	durative, telic	*He walked to the store.*
	Activities:	durative, atelic	*She ran in the park.*
	Achievements:	instantaneous, telic	*She crossed the line.*
	Semelfactives:	instantaneous, atelic	*He sneezed once.*

Table 1: English situation types

Smith (1997), demonstrates that the set of relevant features must be established for each language independently. For example, she argues that there are only three situation types in Navajo: durative events, instantaneous events and statives. These are based on the features grammaticalized in Navajo: dynamic and durative.

[31] See Rothstein (2008) for a contrasting approach to the fundamentals of situation types that argues that Semelfactives are telic.

However, analyses of aspect in particular languages may not relate very directly to the categories and approaches outlined above. Contrast the description of Smith's three Navajo situation types with Axelrod's list of aspect types for Koyukon (Axelrod, 1993).[32] The two languages are both from the Athabaskan language family, a group of languages that are only "shallowly differentiated" (Golla, 2007, p.71). However, Axelrod (1993) argues for three groups of situations in Koyukon with a total of 15 types. These are reproduced in Table 2. These situation types are not defined in terms of the features posited by Smith (1991). Similarly, Mughazy (2005) notes that "there is little agreement in the literature regarding the number of aspectual classes in Egyptian Arabic, as various language-specific classes, such as 'translocatives', 'agentive statives', 'inceptives' and 'pseudo-inchoatives', have been proposed."

State	neuter
	transitional
Motion	momentaneous
	perambulatory
	continuative
	persistive
	reversative
Activity	durative
	consecutive
	repetitive
	directive-rep.
	semelfactive
	bisective
	conclusive
	onomatopoetic

Table 2: Koyukon situation types

Smith (1991, p.10) states that "situation types have no single grammatical marker". Situation type is generally assumed to be a property of the verb phrase or of the sentence/proposition (cf. Krifka, 1989; Smith, 1997; Rothstein, 2008). Situation types are compositionally derived and often dependent on qualities of verbal arguments, sentential adjuncts, inherent properties of the verb or discourse context. In addition to crosslinguistic variation in which situation aspect features, and by extension situation types, are grammaticalized, there is also significant variation in what contributes to the composition of situation type. In other words, languages vary with respect to how much of situation aspect is overtly lexical or morphological, is determined by arguments or temporal adverbials, or is open to inference based on context.

[32]These "aspects" are not defined by Axelrod in terms of the bi-dimensional viewpoint/situation model. However, they do represent a system distinct from that which marks (im)perfective aspect and are described in Axelrod (1993, p.33) as "characteriz[ing] the *manner* of the action", i.e., describing properties of situations.

Many languages provide examples of lexically conveyed features. Stativity is a feature commonly associated with lexical items. English, for example, has a class of stative verbs as well as small classes of verbs that are inherently telic, e.g., *kill*, or atelic, e.g., *beat*. In some languages, features of situation aspect may be marked through overt morphology. For example, according to Rice (2000), Athabaskan languages morphologically mark situation aspect. Zeitoun & Huang (2000) argue that the morpheme *ka-* marks stativity in many Formosan languages. In Russian, a suffix *-nu* marks verbs as members of a Semelfactive class (Comrie, 1976, p.43).

In many languages, verbal arguments play an important role in determining telicity. Telicity is variously described as the property of a situation having an aim, goal, bound, inherent endpoint or change of state. Telicity can be directly affected by properties of arguments, especially the direct object, in many languages, including English. In example (15b), the bounded (quantized) nature of the definite noun phrase (*the apple*) provides an inherent endpoint for the eating event. Similarly, in example (16b), the choice of preposition determines whether the prepositional phrase functions as a destination, again, describing an event with an inherent endpoint or goal, i.e., telic. The examples in (15a) and (16a), however describe events with no inherent endpoint, i.e., atelic.

(15) a. She ate apples.
b. She ate the apple.

(16) a. She walked in the park.
b. She walked to the park.

Finnish provides an example of compositional situation aspect dependent on both inherent properties of lexical items (verbs) and properties of arguments (direct objects). According to Kiparsky (1998) many, perhaps most, Finnish verbs fall into one of two classes, either bounded or unbounded.[33] For the remaining transitive verbs, the boundedness of the situation is determined by the case of the verb's direct object. This sensitivity to case is illustrated in example (17), from Kiparsky (1998, p.2-3). Partitive case (PART) conveys an unbounded situation in example (17a) whereas in example (17b) an object with accusative case (ACC) conveys a bounded situation.

(17) a. Ammu-i-n karhu-a
shoot-PST-1SG bear-PART
'I shot at the (a) bear.' [fin]
b. Ammu-i-n karhu-n
shoot-PST-1SG bear-ACC
'I shot the (a) bear.' [fin]

[33]Kiparsky distinguishes this contrast from the similar category, telicity.

While arguments contribute to the composition of situation type, they are not necessarily definitive. In English, for example, the interpretation of a situation type is sensitive to adverbials or even to discourse context. Consider the sentence, in example (18a). The default interpretation of this sentence is as an Accomplishment (telic), describing a situation with an inherent endpoint established by *the book*, cf. example (18b) which describes an Activity (atelic). However, the closely related sentence in example (18c) describes, again, an Activity. The argument has not changed with the addition of the adverbial although *the book* does not represent an inherent endpoint for the reading event. In fact, *She read the book* can be interpreted as either telic or atelic as is clear from the acceptability of either *in an hour* or *for an hour* in example (18d). These adverbials are commonly used to test for (a)telicity (see Dowty, 1979); note the unacceptability of the atelic situation in example (18b) when paired with *in an hour* in example (18e). With the proper context, both the telic (default) or atelic interpretations of the sentence in example (18a) are, in fact, available.

(18) a. She read the book.
b. She read. (atelic)
c. She read the book for an hour. (atelic)
d. She read the book in an hour / for an hour). (telic / atelic)
e. *She read in an hour.

The interaction with Finnish case above illustrates another significant characteristic of situation aspect already mentioned in §3.2.1: It interacts extensively with other elements of the grammar. For example, according to Dahl (1985, p.28), "[Aspectual] categories tend to be less developed or wholly neutralized in stative contexts." This neutralization is illustrated by the English Progressive which is ungrammatical or has a very specialized meaning with stative verbs. As a further example, also according to Dahl (ibid.), in Beja (Bedawi, iso: bej), stative verb phrases are incompatible with past tense marking.

In summary, situation types can be defined in terms of collections of situation aspect features. Situation types have no single morphological marker but instead are compositional. Situation feature values may also be compositional depending on other sentence elements or context. Features like stativity or telicity, on the other hand, are overtly marked in some languages. Languages vary not only in how situation aspect features are marked but also in what features are relevant. They also demonstrate a high degree of variation regarding the interaction between tense and aspect as well as the interaction with other elements of the grammar.

3.2.3 Analytical variation

The details of tense and aspect systems display large amounts of variation. In addition, while there seems to be some basic agreement about the elements of tense and aspect systems at a fairly high level of abstraction, analyses of tense and (especially) aspect often vary significantly based on differences in linguistic traditions or the linguists' or grammarians' underlying assumptions.

The complexity of tense and aspect often leads to multiple distinct analyses even for the same language. For example, different analyses can result from different assumptions about what is grammaticalized. Specifically, whether the periphrastic expression *être en train de* in French is actually a progressive marker or whether *going to* is a grammatical future tense construction in English, are certainly debatable. This issue becomes quite complex for a language like Turkish, which has a large number of bound morphemes conveying TAM related concepts that, in many languages, are represented by a cline of grammaticalization from lexical verb to inflection. Separating bound morphemes conveying lexical content from those conveying functional (grammatical) content is also complex for linguists working on Athabaskan languages which have notoriously complex fusional morphology (Wilhelm, 2007).

Linguists may also disagree about how semantic categories relate to morphology. For example, consider the following from Bybee et al. (1994, p.126): "Unlike Comrie ... we find it difficult to view the so called present tense as a 'tense', that is, as having to do primarily with deictic temporal reference. What present covers are various types of imperfective situations with the moment of speech as the reference point." There is also a long standing disagreement about whether Russian aspect classes (Perfective and Imperfective) actually convey viewpoint or whether they, in fact, convey situation (lexical) aspect through a feature like telicity or 'boundedness' (Jakobson & Waugh, 1984; Bertinetto & Delfitto, 2000).[34]

In addition to variation in assumptions about grammaticalization and semantic content, there is also considerable variation among analyses based on differences in the underlying models of tense and aspect assumed; linguists disagree about what constitutes aspect, what the fundamental elements are and how the elements interact (Sasse, 2002).[35] Some of this disagreement is undoubtedly related to the history of various linguistic traditions. According to Sasse (2002, p.212-213), historically, the Slavic and Romance traditions restricted the study and analysis of aspect to grammatical (viewpoint) contrasts whereas the Anglo-American tradition focused almost exclusively on characteristics of situations. These contrasting views of aspect were a direct consequence of the aspectual systems in the languages under study.

Current models of aspect, e.g., Smith (1997), tend to view aspect as bi-dimensional with two interrelated components. However, there are also uni-dimensional approaches. These approaches tend to either reject some elements as not central to aspect or amalgamate all aspectual elements (Sasse, 2002, p.213-217). Post-Vendlerian approaches (e.g., Dowty, 1979; Bach, 1981) assume that the possible aspect values are STATE, ACTIVITY, ACCOMPLISHMENT or ACHIEVEMENT and, typically, assume that sentence aspect is compositionally derived from inherent values of the verb which may be "recategorized" through the action of additional material such as noun phrases, prepositional phrases, or aspect morphology (e.g. progressive marking). In uni-dimensional approaches, grammatical bounds are viewed as operators on aspect values rather than as elements of an independent dimension of aspect (see also de Swart, 1998). This approach to aspect is generally applied to English and other Germanic languages with little or no grammatical aspect (Sasse, 2002).

The English Progressive provides the context for a specific example of the broad impact

[34]For many additional references relevant to this issue see Slabakova (2004, p.5, footnote 4).

[35]See Sasse (2002) for a concise overview of the issues.

of the underlying model on the resulting analysis. Smith (1997) and Comrie (1976) both assume that the English Progressive is a type of imperfective viewpoint. In contrast, Vlach (1981), as described in Sasse (2002, p.215), argues that the progressive is a function, or operator, that makes sentences stative. According to this uni-dimensional model of aspect, sentences have no viewpoint aspect value. The only aspect value is the situation type and, in a sentence with a Progressive form, the Progressive changes the situation type of the predicate to stative. Therefore, under this model, all Progressives are stative. This contradicts the analysis of English Progressive by Smith (1997) or Comrie (1976) (among others) wherein the standard English Progressive sentence is not stative, as progressive viewpoint is generally incompatible with stative situations.

Less fundamental variation in the underlying model may also significantly affect aspectual analysis. Consider a contrast in the assumptions by Smith and Comrie: Smith (1997) argues that all sentences, in every language, have viewpoint aspect values, either through an overtly marked contrast or covertly. According to her model (Smith, 1997, p.67), all simple English sentences not marked Progressive convey Perfective aspect. Smith argues that the semantic content conveyed by the Perfective depends on situation type. For example, a telic situation in the Perfective viewpoint is interpreted as completed while atelic Perfectives are interpreted as merely terminated. Comrie (1976, p.25), on the other hand, argues that simple English verb forms are underspecified for aspect, i.e., they can convey perfective or imperfective aspect depending on the context. Therefore, for the English sentence *He ate bananas*, Smith's model analyzes the viewpoint as Perfective while, according to Comrie, this habitual sentence is Imperfective. The underlying assumptions of the linguists lead to very different analyses of the English viewpoint aspect values for the same sentence.

3.3 Summary

Tense and aspect systems are complex, inter-dependent and language-specific. Analyses of these systems are also complex. With aspect, and to a lesser extent with tense, the model assumed by the linguist greatly affects the conclusions drawn about the relations between aspectual elements and the aspectual elements themselves. In a very real sense, every analysis of tense and aspect is a hypothesis or, perhaps more accurately, a series of hypotheses. Implementation within even a fairly small grammar is a good way to test these hypotheses. In §4, I discuss how the meta-modeling approach is designed to assist in this process and lay out the scope of this work.

4 Goals

The main goal of the Grammar Matrix customization system is to support the linguist-user's development of an analysis through grammars that jump-start the engineering process. This means that the code should scale up, providing a basis for the continued development of the grammar and, specifically, it means that a customization should reflect the linguist-user's own analysis. With that in mind, this implementation of tense and aspect strives to

recognize tense and aspect marking, to place reflexes of that marking into the semantics and to constrain ungrammaticality while imposing as little pre-defined analysis as possible on the user. In order to avoid restricting the user to a specific model of tense and aspect, this implementation uses a meta-modeling approach. I discuss this meta-modeling approach in §4.1 and the scope of the implementation in §4.2.

4.1 Meta-Modeling

The meta-modeling approach consists of building an implementation that provides a means for the user to define grammar components consistent with their own tense and aspect analysis instead of offering a collection of pre-defined analyses.[36] This approach is primarily motivated by the complexity and language-specific nature of tense and aspect systems and the need to accommodate variable assumptions and models.[37] Without common agreement about terminology, without predictable hierarchies, without a universally applicable set of building blocks in a commonly held model, it is difficult to develop a useful set of pre-defined analyses. Attempting to model tense and, especially, aspect strictly in terms of a finite number of developer-defined analyses, would require some very limiting choices. Users would have to mold their analysis to match the one(s) presented by the customization. This works against the goal of assisting the linguist. What is undertaken instead, using this meta-modeling approach, is the development of a limited amount of structure within which the user defines their analysis using their own terminology and assumptions.

The implementation is discussed at length in §5 however, here, I illustrate the meta-modeling approach with just a few details. The implementation provides a tense feature and two aspect related features. These features, TENSE, ASPECT and SITUATION, are defined as features of a neo-Davidsonian event variable (see Parsons, 1990). The event variable is standard for Matrix grammars, as is this location for the definition of event-related semantic features, following the approach taken by MRS (Copestake et al., 2005) and applied in the English Resource Grammar (ERG) (Flickinger, 2000). These features are pre-defined in the customization system. This provides a framework within which the user provides the possible values for each feature and the relationships between the features as defined in hierarchies. The user also determines the use made of these features in the broader analysis.

The two pre-defined aspect features, ASPECT and SITUATION, reflect the bi-dimensional approach to aspect described in §3. However, other than these names there is no other distinction between the features themselves and there is no requirement that both of these features be used; a uni-dimensional approach or a partial analysis might make use of only one aspect feature. This neutral approach accommodates various underlying models of aspect. On the other hand, there are some underlying assumptions built into this implementation

[36] The term *meta-modeling* intends to evoke the idea of a higher level of structure; meta-modeling is to the building of grammar models as meta-data is to the organization of data.

[37] Of course, such complexity is not restricted to tense and aspect, but tense and aspect form a particularly notable case and were also the first phenomenon encountered in the development of the Grammar Matrix customization system that had such requirements. The meta-modeling techniques were subsequently adopted in other parts of the system as well, including Drellishak's analysis of gender and number (Drellishak, 2009).

that do impact or restrict the definable analyses. Tense, viewpoint and situation features are assumed to be semantic, not syntactic, consistent with the conclusions in Kibort (2008).[38] A basic assumption of MRS, also reflected in the Matrix, is that the elements of the semantic representations must be monotonically composed, i.e., semantic values cannot be removed or changed but instead can merely be appended or made more specific. This lends itself to semantic analyses based on underspecification. This requirement of monotonic composition makes analyses of tense and aspect based on operators that change aspect values, as in de Swart (1998), incompatible with Matrix grammars. In addition, the Matrix grammars are developed with an HPSG-based lexicalist bias. This translates into a preference for proliferating lexical types and lexical rules over phrase-structure rules. Therefore, this implementation concentrates on providing flexibility with respect to those lexical types and rules as opposed to construction-based solutions.

While the meta-modeling approach cannot accommodate all analyses, it does provide a high level of flexibility to the user, allowing them to define a tense and aspect system that reflects their own analysis. In addition to accelerating the development of a grammar, this flexibility also allows the linguist-user to do comparative implementations of various options relatively easily. This makes the tense and aspect customization a valuable tool for hypothesis testing as well as for jump-starting a broad-coverage grammar. Given the complex nature of tense and aspect, this may be a significant additional benefit of this meta-modeling approach.

From a practical perspective, being able to quickly define and test a grammar model can potentially save the grammar-engineer a lot of time by revealing issues and forestalling dead end approaches. The more interesting question may be whether linguists who are not grammar engineers might also find this a useful tool for testing hypotheses about tense and aspect. Implemented grammars provide significant benefits for hypothesis testing (see Bender, 2008, for discussion). While the grammars currently produced by the customization system are small they are sufficient to test interactions between many different phenomena related to tense and aspect. If the phenomena interacting in an analysis are among those definable through the questionnaire then there is no need for computer programming skills to create a model of the analysis; the process would entail defining the relevant phenomena on the questionnaire and requesting a grammar, testing this grammar by parsing grammatical and ungrammatical sentences with the customized grammar and examining the parse results for expected or unexpected behaviors.

In §4.2, I move on to discuss specifically what elements of tense and aspect are covered or, more precisely, directly addressed in this implementation. What is actually covered by a grammar customized through this implementation is dependent on the analysis of tense and aspect developed by the user.

[38] This is a working hypothesis and, as pointed out to me by Dan Flickinger (p.c.), potentially falsifiable.

4.2 Scope

This implementation of tense and aspect in the Matrix customization system is designed to provide a platform for defining basic tense and aspect systems including semantic categories, morphological marking and syntactic constraints. It facilitates the definition of language-specific analyses. This implementation accommodates a useful array of prominent phenomena associated with tense and aspect including:

- absolute tense
- viewpoint aspect
- properties of situation aspect
- inflectional morphology
- fusional morphemes
- auxiliary verb constructions
- auxiliary verb plus inflection combinations
- inherent lexical properties
- verb classes
- constraints on auxiliary complements
- constraints on inflectional attachment
- language-specific variation, for example:
  - traditional terminology
  - tense-related remoteness distinctions
  - non-Indo-European aspect systems

In §5, I present details of this implementation and illustrate its utility and flexibility. However, before moving on, there are a few important things to note about the context assumed for the grammars produced by the customization system. Some phenomena, such as discourse context, are completely outside of the scope of Matrix grammars. Matrix grammars parse strings/sentences without any context. Since situation type is, at least in part, dependent on discourse context, the calculation of situation type cannot be accomplished within the grammar. This implementation, therefore, provides no specific support for situation aspect composition and assumes it occurs post-syntactically. Semantic inference is also out of the scope of Matrix grammars. The MRS (semantic) representations generated by these grammars are assumed to act as input to external semantic post-processing. This supposition of a post-syntactic process is consistent with the observation that elements of temporal interpretation are dependent on discourse context. It is also consistent with the fact that

the semantic categories are language-specific and as such do not represent an inter-lingua.[39] In addition, morphological and phonological analysis also falls outside of the scope of the Matrix grammars currently produced. Matrix grammars assume an external pre-processing step so input to the grammars consists of sequences of regularized morphemes.

5 Details

5.1 Process

From the point of view of the linguist-user, the Grammar Matrix customization system consists of a web-based questionnaire[40] containing a myriad of questions related to a range of language phenomena. Answers to the questions populate a "choices" file consisting of key-value pairs. Based on these choices, a small, working, Matrix-based HPSG grammar is created and made available for download.

From the perspective of the developers, the system includes:

- the Matrix core grammar
- code that creates the HTML questionnaire
- Python code that validates the choices on the questionnaire ensuring that they are internally consistent and, to a certain extent, linguistically reasonable
- Python code that processes the key-value pairs obtained from the questionnaire
- Python code that generates a TDL file that defines the language-specific elements of the grammar as determined by the choices
- code that wraps up and compresses all of the relevant files for the grammar and its parser/generator
- Python and Lisp code that creates and runs test-suites for regression testing purposes

Even within the code that specifically writes the TDL files there is a significant amount of infrastructure code not specific to particular phenomena. This code has been contributed by the various members of the Grammar Matrix development group, especially Scott Drellishak who wrote, in addition to several libraries and other bits of the current infrastructure, the code that supports 'iterators' or parts of the questionnaire where the user can enter unbounded items of some type, Kelly O'Hara who developed code for handling the attachment, ordering and interaction of affixes (O'Hara, 2008), Antske Fokkens who is currently expanding the word order coverage and Emily Bender whose fingerprints are everywhere. The Grammar Matrix development philosophy encourages generalizing code as well as code extension and reuse. This creates a more consistent and more sustainable system.

[39] Matrix grammars have been used in various machine translation experiments and as such the inter-lingual possibilities of MRS representations are of interest.

[40] The Grammar Matrix customization system questionnaire is available at: `http://www.delph-in.net/matrix/customize/matrix.cgi`

The process of implementing a new library for the Grammar Matrix customization system usually entails basically four tasks:

- researching and developing analyses of the phenomenon,
- developing the content and format of the questionnaire,
- writing rules for validating the questionnaire to ensure that the obligatory parts are complete and reasonable, and
- writing code that will generate an appropriate analysis in TDL given the questionnaire choices.

To present the results of this process for tense and aspect, I will first give an overview of the parts of the questionnaire relevant to tense and aspect, walking through each relevant section of the questionnaire and discussing how choices made on the questionnaire are converted to language-specific TDL in the output grammar.[41] Both tense and aspect are optional on the questionnaire; users are free to build starter-grammars with tense or aspect or neither. Questions related to tense and aspect are mainly in three sections of the questionnaire: Word Order, Tense and Aspect Features and Lexicon.

5.2 Word Order

Word order relates to every other phenomenon in a language. Since there are word order issues that specifically affect the implementation of auxiliaries, on the Word Order section of the questionnaire, users must provide information about the existence and behavior of auxiliaries:

- Does your language have auxiliaries?
- What order do they occur in with respect to their complements?
- What kind of complements do they take (V, VP, S)?
- Additional questions related to free word order languages and verb clusters

While users may optionally not define tense and aspect through the questionnaire, users are required to declare on the Word Order page whether or not their language has auxiliaries. Although auxiliary verb constructions, more often than not, convey tense and/or aspect information, auxiliaries may play a role in several other phenomena including negation and question formation. If a language has auxiliaries at least one auxiliary verb type must be defined in the Lexicon section. If a language has auxiliaries there are also consequences with respect to the FORM feature discussed in §5.3.4.

In the Matrix customization system, auxiliaries are assumed to be verbal heads that take complements. In the current version of the customization system, auxiliary-complement

[41] Parts of this discussion involve interactions with other libraries or modification of code that was extant when I began this implementation.

order is further assumed to be fixed for all auxiliaries in a given language.[42] The current implementation of word order assumes that there are three possible auxiliary complement types: verb (V), verb phrase (VP) or sentence (S).[43]

The main questions related to tense and aspect are on the Tense and Aspect Features page and in the verbal and inflectional sections of the Lexicon page. These are discussed below. In addition, use of the Other Features page is described as it applies to defining verb classes. The sections are discussed in the order they appear on the questionnaire.

5.3 Tense and Aspect Features

The Tense and Aspect Features page is divided into two sections wherein the linguist-user may define values of the semantic features TENSE, ASPECT and SITUATION and the purely syntactic feature FORM. These values are reflected as type hierarchies in the customized output grammar. Once defined, these features become available in later sections of the questionnaire. For example, values provided here become available for defining inflectional elements and auxiliaries and for constraining auxiliary complements. I will discuss each of these features in turn below.

5.3.1 TENSE

The questionnaire asks the linguist-user to define a hierarchy of possible values for the feature TENSE. There are two methods provided for defining the TENSE hierarchy. One allows the user to select among the typologically common tense values: past, present, future, non-past and non-future. For each given value, it is also possible to augment the hierarchy to accommodate a single layer of subtypes. This layer is primarily designed to accommodate remoteness distinctions. For example, according to Comrie (1985, p. 90),[44] the Niger-Congo language, Haya (iso: hay), makes three past, and two future, remoteness distinctions. Figure 4 is a screenshot of the portion of the questionnaire that selects 'past' as a value in the tense hierarchy and defines its subtypes, per Comrie's description.

The second option provided for defining a TENSE hierarchy is to "Build your own hierarchy". This option allows the linguist complete flexibility with regard to the shape of the hierarchy and the value labels. This is useful if, for example, the linguist desires different value labels. For example, Comrie (1985, p. 88) suggests that the discontinuous (cyclical) tense system of the Australian Aboriginal language, Burera (a.k.a. Burarra, iso: bvr), might be well-described as having 'close' and 'remote' tenses as opposed to 'past' and 'present' since events occurring yesterday or now are conveyed by the 'close' form while those occurring earlier today or before yesterday are conveyed by the 'remote' form. The process to build-your-own hierarchy is to enter a value and then choose one or more supertypes for that value from a drop-down list. Figure 5 shows the questionnaire entries for producing a simple

[42] Whether there is evidence of languages with variable auxiliary complement order is currently being investigated in connection with the continuing development of the word order library by Antske Fokkens.

[43] The existence of languages with auxiliaries taking sentential complements is speculative at this point.

[44] Comrie notes that the data is from the unpublished work of Ernest R. Byarushengo.

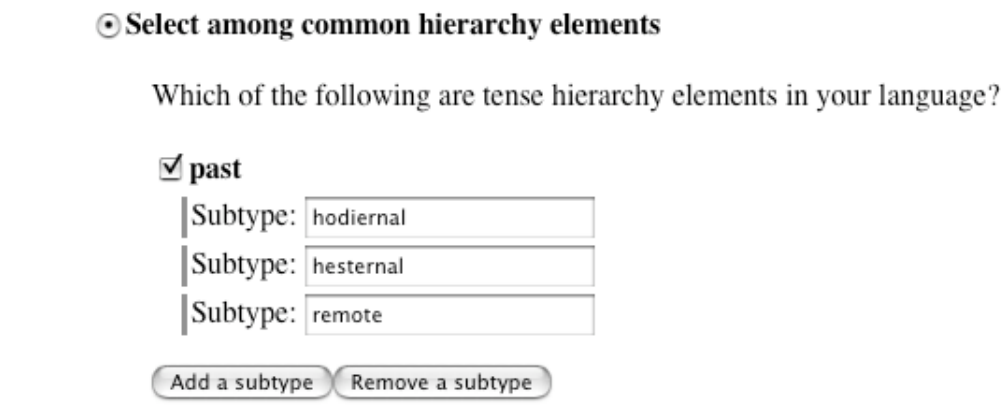


Figure 4: Questionnaire: adding remoteness distinctions

hierarchy of two tense values such as described above for Burera. Example (19) provides the corresponding (fairly uninteresting) type definitions from the output grammar. The values *close* and *remote* are defined as subtypes of TENSE.

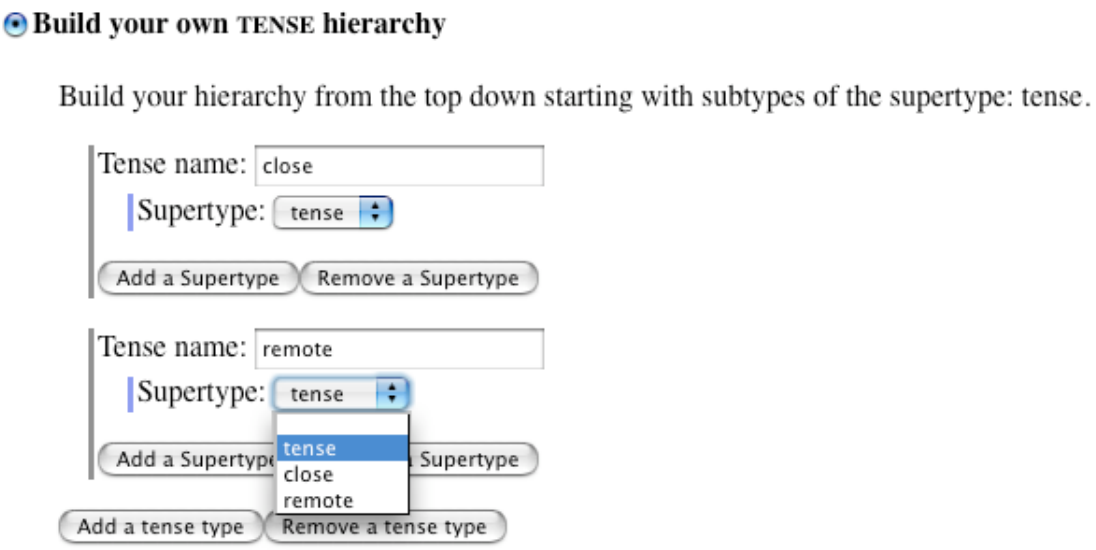


Figure 5: Questionnaire: building a tense hierarchy

(19)
```
close := tense.
remote := tense.
```

The build-your-own option is also useful for defining tense systems requiring additional levels in the hierarchy. For example, the Amerindian language, Kiksht (iso: cch) (Comrie, 1985), has four grammaticalized past distinctions and at least two of those distinctions are further subdivided into 'early' and 'late'. Table 3 contains data from Hymes (1975), as

described in Comrie (1985, p.100). To the data, I have added the terms in brackets which correspond to the nodes in the tree in Figure 6.[45] This hierarchy represents an analysis of this data that defines a deeper hierarchy than can be created by augmenting the hierarchy of a selected value. With this deeper hierarchy, the intermediate tense values, e.g., remote, can serve as partially underspecified values.[46]

ga(l)...u-	remote past	[early-remote]
ga(l)...t-	from one to ten years ago	[late-remote]
ni(g)...u-	from a week to a year ago	[early-far]
ni(g)...t-	last week	[late-far]
na(l)-	yesterday or preceding couple of days	[middle]
i(g)-	earlier today:	[near]
	?i(g)...u- earlier on today, but not just now	
	?i(g)...t- just now	

Table 3: Kiksht past distinctions

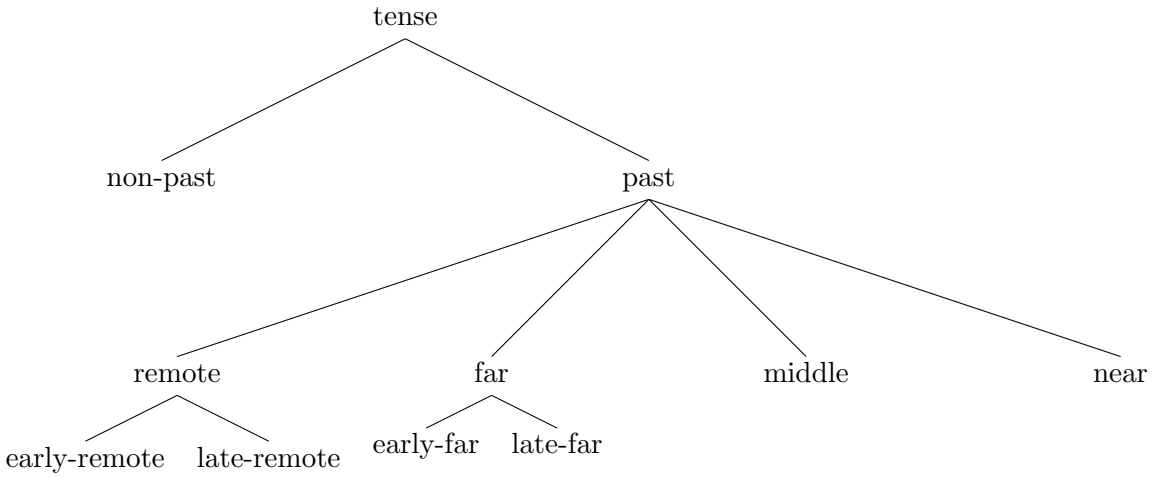


Figure 6: Kiksht (past) tense hierarchy

The feature TENSE is defined in the Matrix core grammar as one of a collection of features of the semantic event variable. Specifically, the event variable, typically itself the value of the feature INDEX, has a feature (E) which takes as its value a feature structure of type *tam*. The type *tam*, as defined in the Matrix core, has the features TENSE, ASPECT and MOOD as attributes.[47] Values of TENSE, as well as those of the features ASPECT and MOOD, are defined as type *sort*, i.e., types with no internal features.

[45]Comrie expresses uncertainty about the u-/t- distinction in the final set in Table 3 as indicated in the table by the question marks. I have left that distinction out of the tense hierarchy tree as it is not relevant to the point at hand.

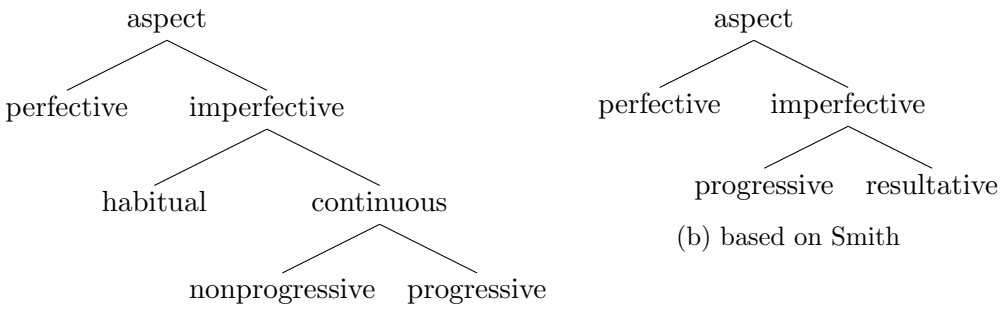
[46]This is by no means the only way this data could be structured.

[47]Since some languages appear to have no grammaticalized tense, there has been some discussion of moving

5.3.2 ASPECT

The questionnaire essentially assumes the bipartite division in aspect discussed in §3 and provides two separate features of the event variable related to aspect: ASPECT and SITUATION.[48] SITUATION will be discussed in §5.3.3. ASPECT is provided to accommodate grammatical aspect categories and, like TENSE, is already defined in the Matrix core. As with TENSE, the feature ASPECT ranges over a set of terms defined by the linguist-user and arranged into a type hierarchy so as to allow for partial underspecification.

The questionnaire allows the linguist-user to build an aspect hierarchy following the same build-your-own process provided for TENSE. This allows the linguist control over the shape and content of the hierarchy. The trees in Figure 7 illustrate two contrasting analyses of viewpoint aspect based on the hierarchies listed in Comrie (1976, p.25) and inferred from the text in Smith (1991, p.219-227). These include the aspect categories discussed briefly in §3.2. Both of these hierarchies can be easily defined through the questionnaire. Figure (8) lists the two corresponding sets of TDL code that the customization system produces for these definitions.



(a) based on Comrie

(b) based on Smith

Figure 7: Contrastive English aspect hierarchies

5.3.3 SITUATION

The SITUATION feature is primarily useful for representing inherent lexical qualities or those conveyed through overt morphological marking.[49] The questionnaire allows the linguist-

at least the tense entry in the AVM out of Matrix core which is specifically designed to contain only those aspects of a grammar that are universally applicable. If removed from Matrix core, the feature would be added to individual grammars as needed. Moving TENSE out of Matrix core would make no difference in the functioning of the grammars but would be more consistent. It has not yet been moved out due to issues of compatibility with previous matrix grammars.

[48] If both features are used, in a bi-dimensional analysis, ASPECT would correspond to viewpoint aspect and SITUATION to situation aspect. A uni-dimensional analysis (cf. §3.2.3) would use only ASPECT.

[49] This may more accurately reflect uses of the term 'lexical aspect'. Therefore, it might seem clearer to use a feature name like 'lexical', as opposed to 'situation'; however, given the lexical approach assumed in these grammars, there are already many 'lex'-coinages within the grammars so I deemed any 'lex'-based

```
perfective := aspect.
imperfective := aspect.
habitual := imperfective.
continuous := imperfective.
nonprogresssive := continuous.
progressive := continuous.
```

(a) TDL: based on Comrie hierarchy

```
perfective := aspect.
imperfective := aspect.
progressive := imperfective.
resultative := imperfective.
```

(b) TDL: based on Smith hierarchy

Figure 8: TDL: generated from contrastive hierarchy definitions

user to create a hierarchy of values for the feature SITUATION. These values then become available for defining verb types and constraining auxiliary complements. SITUATION has been added as a feature by this implementation. As such, it is not included as a feature of the event variable in the Matrix core. Therefore, if SITUATION values are defined on the questionnaire, the feature SITUATION is added as another feature on the *tam* matrix. Then the categories provided through the questionnaire are defined as values of SITUATION. Example (20) contains the TDL generated in a customized grammar defined through the questionnaire to include a SITUATION feature and a very simple hierarchy of *situation* type values.[50] Figure 9 contains a portion of the feature structure that results from parsing the simple non-stative, past tense sentence, *the cat chased the dog* with this grammar.

(20)
```
situation := sort.
tam :+ [ SITUATION situation ].

stative := situation.
nonstative := situation.
```

$$\begin{bmatrix} \text{SYNSEM.LOCAL.CONT.HOOK.INDEX} & \begin{bmatrix} \textit{event} & \\ \text{E} & \begin{bmatrix} \textit{tam} & \\ \text{TENSE} & \textit{past} \\ \text{ASPECT} & \textit{aspect} \\ \text{MOOD} & \textit{mood} \\ \text{SITUATION} & \textit{nonstative} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 9: Feature structure extract (simplified): SITUATION added to the event feature (E)

name too confusable.

[50]This simple example ignores issues of mood and assumes that viewpoint aspect is underspecified in this context.

It is worth noting that the only real restrictions on the values of ASPECT or SITUATION are that they be of type *sort*.[51] It is therefore possible to use the properties of dynamicity, durativity, etc. as the basis for a hierarchy of situation aspect types.[52]

5.3.4 FORM

In many languages, the marking of tense and/or aspect involves a system of auxiliaries. The feature FORM is used to capture the distribution of different inflected forms of main verbs with and without auxiliaries. Specifically, we assume that if a language has auxiliary verbs, it also groups inflected verb forms in to finite and non-finite classes.[53] The list of non-finite English verb forms includes infinitive, past participle and present participle. We assume that finite verbs, canonically, can head stand-alone clauses, while non-finite verbs (canonically) cannot. The grammars produced by the customization system distinguish forms of the verb through the syntactic HEAD feature, FORM. This feature is a variation on the VFORM feature dating back to at least Pollard & Sag (1994). Figure 10 contains a portion of the feature structure that results from parsing the simple finite sentence, *the cat chased the dog*.

$$\begin{bmatrix} \text{SYNSEM.LOCAL.CAT.HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{AUX} & \text{-} \\ \text{FORM} & \textit{finite} \end{bmatrix} \end{bmatrix}$$

Figure 10: Feature structure extract (simplified): FORM as a HEAD feature

If the choice made on the questionnaire indicates that the language has auxiliaries the grammar produced will contain *finite* and *nonfinite* values for the feature FORM. In addition, every grammar has a root condition, i.e., a condition that must be met for a string to constitute a complete sentence. If languages make a finite/non-finite distinction the root condition is set to state that sentences must be headed by a verb of FORM *finite*.

Alternatively, if the language has no auxiliaries but does make a syntactically relevant distinction, the questionnaire provides the option to, nevertheless, define finite and non-finite FORM values. Depending on the language, this might be useful for specific forms associated with imperatives or subordinate clauses.[54]

[51] That said, there are general typographical restrictions on entries for the entire questionnaire disallowing reserved names, restricting special characters and prohibiting blank space within entries.

[52] The current implementation of the questionnaire requires a hierarchy of types approach as opposed to a binary feature based approach. While the descriptive power of the two approaches is equivalent, providing the binary approach as an option is under consideration.

[53] This may well be an incorrect assumption. If it is found to be too broad a generalization then specific questions about the existence of a finite/non-finite distinction will be added to the questionnaire. However, the consequence of being wrong in this case is the inclusion of a bit of unused code in the grammar. While potentially irritating, this will not affect the behavior of the grammar.

[54] While the customization system does not yet cover these phenomena, the linguist-user might want to associate the appropriate forms with the corresponding inflectional morphology through the questionnaire.

In this section of the questionnaire, the linguist-user can also add subtypes of *finite* and *nonfinite*. In many languages with auxiliaries, non-finite form subtypes are useful for constraining auxiliary complement forms. Figure 11 shows the addition of non-finite forms for English on the questionnaire. These FORM values are translated into type statements in the customized grammar which define the form hierarchy in Figure 12.

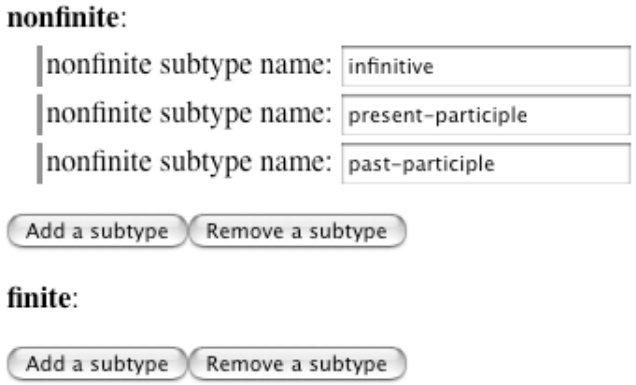


Figure 11: Questionnaire: adding non-finite forms

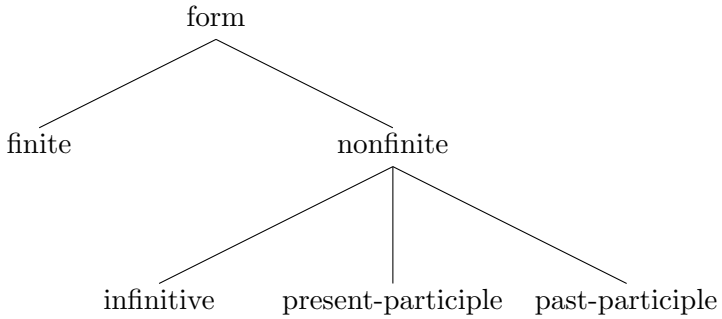


Figure 12: Extended FORM hierarchy

TENSE, ASPECT, SITUATION and FORM feature values can be defined on the Tense and Aspect Features page. These features then become available on the Lexicon page for association with lexical items or inflectional morphemes.

5.4 Other Features

In addition to the sections of the questionnaire that are directly related to tense and aspect, there is another section of the questionnaire that may be useful for the linguist-user defining a tense and aspect system through the questionnaire. The Other Features page was developed as a part of the case and agreement work described in Drellishak (2009). Broadly, this section allows the linguist-user to define some features not otherwise definable on the questionnaire. Relevant to tense and aspect is the user's ability to define syntactic HEAD features, each with

a corresponding hierarchy of values (types). Once defined, these features and values become available for use in the Lexicon section.

This functionality can be useful for defining certain kinds of verb classes. Verb classes may be semantically based but they often consist of like elements with no semantic, or only quasi-semantic, characteristics, i.e., the classes have become grammaticalized or are not clearly associated with a specific semantic feature. For example, in many Romance languages, verbs fall into auxiliary-selection classes which are generally analyzed as syntactic, i.e., not semantically based. Specifically, French verbs fall into one of two classes with respect to auxiliaries. The verb *parler* is compatible with the auxiliary verb *avoir* as in example (21a), but not *être*, as in (21b). The ungrammaticality of (21b) contrasts with the grammatical example in (21c) since the verb *arriver* is a member of the class of verbs compatible with *être*.[55]

(21) a. Il a parl-é.
he AUX(avoir) speak-PST
'He spoke.' [fra]

b. *Il est parl-é.
he AUX(être) speak-PST
'He spoke.' [fra]

c. Il est arriv-é.
he AUX(être) arrive-PST
'He came.' [fra]

A syntactic feature can be used to distinguish members of the auxiliary-selection classes. Such a feature can be defined on the Other Features page. Values of the feature can then be used in the Lexicon section of the questionnaire as discriminators when defining two classes of verbs, one for each auxiliary.

5.5 Lexicon

Once features have been defined, they can be associated with sentential elements, i.e., with lexical items or inflection.[56] Elements of aspect, for example, can be inherent properties of lexical verbs. Stativity is a property of certain English verbs and perfectivity is a property of the class of derivationally constructed Russian Perfective verbs. In Matrix grammars, inherent properties of verbs are established primarily through lexical-type definitions; each lexical item inherits from a lexical type, inheriting all of the properties of the type. Lexical type definitions can also be used to specify properties of arguments. For example, auxiliary verbs may specify requirements on the form as well as various tense and aspect properties of its complement.

[55] I have ignored aspect in the glosses as it is not relevant to the choice of auxiliary in French.

[56] Of course, features may be introduced through the construction itself, i.e., through specific versions of head-complement rules. However, as we are assuming a lexicalist approach, we do not provide that option though the questionnaire.

In Matrix grammars, inflection is handled through lexical rules. Feature values associated with inflection are defined in these rules. Lexical rules can specify values on the inflected word or any of the word's arguments.

On the Lexicon page, lexical types are defined and stems (uninflected lexical entries) are associated with those types. In addition, inflectional rules are defined, attachment of the inflectional morphology is constrained and features are associated with lexical rules. In §5.5.1 and §5.5.2 below, I will describe the definition of (main) verbs and auxiliaries as it relates to tense and aspect. Then in §5.5.3, I will discuss briefly how these types, and the defined features, interact with the customization system's morphological infrastructure.

5.5.1 Verb types

Any number of verb types can be defined on the questionnaire. For each, an optional type name may be entered,[57] argument structure must be specified and any number of verb stems, along with their associated semantic predicate strings, may be entered. Feature-value pairs may be chosen from lists of those previously defined on the questionnaire, including those for tense and aspect. These features may then be specified on the verb or on its arguments (subject or object).[58]

Abstractly, lexical types provide a means of organizing the lexicon. Lexical entries inherit from types and so inherit the features of the type. This removes redundancy from the lexicon. Defining a verb type therefore creates a partial definition of each verb of that type. For example, most Russian verbs are members of a pair of related verbs, one Perfective and one Imperfective.[59] This can be modeled with separate verb types: a Perfective type and an Imperfective type. Figure 13 shows the definition of Perfective (transitive) and Imperfective (transitive) verb types on the questionnaire as well as a couple of verb stems for each type: *sjed* (*to eat* – Perfective), *prochitaj* (*to read* – Perfective), *jed* (*to eat* – Imperfective), *chitaj* (*to read* – Imperfective).[60]

Example (22) contains an example of the TDL type definitions created by the definition of the Russian Perfective verb type. This type (Perfective-verb-lex) inherits from the transitive lexical type and also contributes a value for ASPECT. The stems defined on the questionnaire for this verb type inherit from Perfective-verb-lex.

(22)
```
Perfective-verb-lex := transitive-verb-lex &
      [ SYNSEM.LOCAL.CONT.HOOK.INDEX.E.ASPECT perfective].
```

Note that this analysis assumes that distinctions in viewpoint differentiate these verb classes. As mentioned in §3.2.3, alternative analyses asserts that the distinguishing characteristic of 'Perfective' verbs is one of telicity (or 'boundedness'), a property of situation aspect, not perfectivity. If an appropriate type hierarchy is defined for situation aspect the

[57] If no type name is provided a generic name is created.

[58] The current version of the customization system offers only simple transitive or intransitive valence options for non-auxiliary verbs.

[59] This is true, in fact, of most Slavic languages.

[60] These spellings are orthographic transliterations of the Russian.

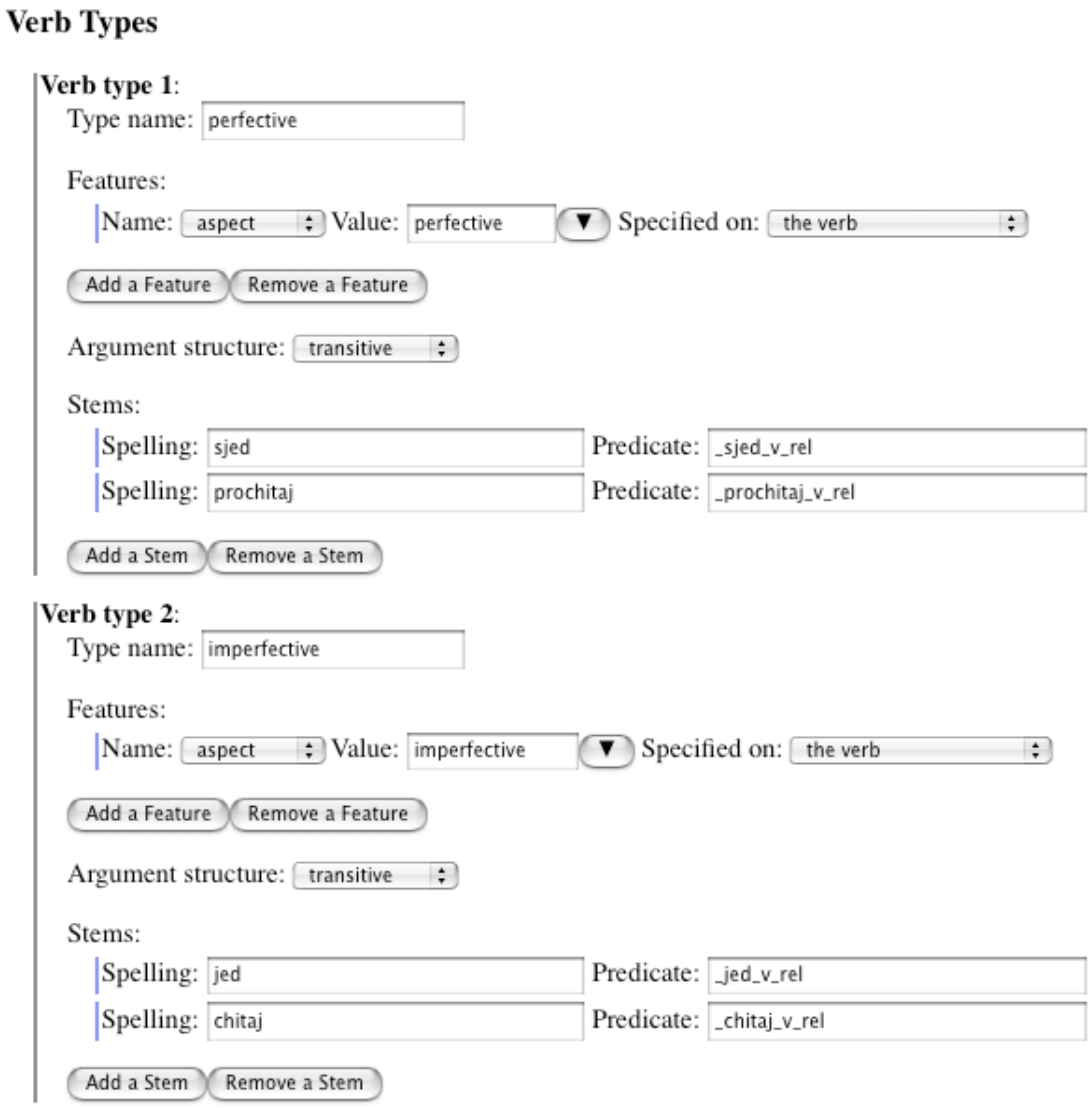


Figure 13: Questionnaire: defining Russian verb types

feature-value pair [SITUATION *telic*] becomes available as an alternative for defining a Perfective verb class. A third approach might be to define syntactic verb classes. A syntactic feature and its possible values can be defined on the 'Other Features' page. Assuming an appropriate feature and its possible values have been defined, values can be selected from the provided list and attributed to the verb although, under this analysis, no reflex of the contrast ends up in the MRS — a, perhaps, undesirable outcome. Any of these analyses is possible to define through the questionnaire; the choice depends on the user's analysis.

In many languages, there are grammaticality distinctions related to verb classes. In particular, the aspect associated with particular verb classes can affect its interaction with other aspects. For example, (23) illustrates the stative/non-stative interaction with progressive aspect in English. Stative verbs, like *understand*, are not compatible with the standard

progressive (be + ing) construction.[61]

(23) a. She was eat-ing the pie.
she AUX.PST eat-PROG the pie [eng]

b. *She was understand-ing the answer.
she AUX.PST understand-PROG the answer [eng]

This can be modeled as a constraint on the complement of the progressive auxiliary. On the questionnaire this entails first defining a class of stative and non-stative verbs much as perfective and imperfective classes were defined for Russian (see Figure 13). Defining the constraint on the auxiliary is discussed below in §5.5.2. The definition of English stative and non-stative feature values, as well as verb types and constraints on the English Progressive based on those features, are illustrated in §5.6.

5.5.2 Auxiliary verb types

Auxiliary is a category on a continuum with serial verb constructions on one end and inflection on the other (Anderson, 2006). Whether a specific element in a given language is an auxiliary is open to interpretation. In the Grammar Matrix customization we define an auxiliary as an element that functions as the head of a verbal projection, takes a V (verb), VP (verb phrase) or S (sentence) as its complement and (canonically) contributes tense or aspect information, mood/modality or negation.

On the questionnaire, auxiliary types are defined separately from other verb types. There are two distinct parts to the definition of an auxiliary verb: describing the auxiliary itself and describing its complement. As with regular verb types, the user enters an auxiliary type's name and stems associated with the type.[62] However, not all auxiliaries introduce their own semantic predicates. For example, assuming a particular analysis of English, auxiliary *be* contributes nothing but constraints on the tense and aspect of the event described by the main verb while the modal *can* contributes a semantic notion of ability or possibility to the phrase.[63] Therefore, one of the choices when defining auxiliary types on the questionnaire is whether the auxiliary contributes a predicate.

If an auxiliary introduces its own predicate then it also introduces its own event feature on which features of the auxiliary can be defined. While the advantages of declaring tense and aspect related features on the auxiliary, as opposed to defining them as constraints on the complement, is not clear, it is a logical possibility. The questionnaire allows the linguist-user to associate features with an auxiliary type by choosing features and values from a list of those previously defined. If the auxiliary does not introduce a predicate then the

[61]There are of course examples of apparent progressive-stative combinations in English such as 'I am loving this color!' but this usage is restricted to small set of stative verbs and entails a significant meaning alteration so it is not generally given the same analysis, e.g., see Bertinetto (2000, p.560).

[62]While any number of stems are allowed, it is likely that in many cases auxiliary types will not include more than one stem per type.

[63]See Palmer (1986) for a discussion of modality.

INDEX feature is effectively identified with the INDEX of the head of its complement, i.e., the semantics of the auxiliary and that of the main verb are identical.[64]

General options about type and ordering of auxiliary-complements are defined on the Word Order page, as discussed in §5.2 above. Here, in the Lexicon section of the questionnaire, the user further describes the complement of each auxiliary type. Feature-value pairs previously defined on the questionnaire can be selected from drop-down lists to specify values of the complement. A value for the FORM feature is obligatory; it constrains the form the complement verb must take.[65] In the customized grammar, the selected feature-value pairs translate to specifications of values for features of the head (verb) of the auxiliary's complement. Returning to the example of the English progressive, from §5.5.1: Given appropriate feature definitions, the feature SITUATION with value *nonstative* can be selected as a constraint on the complement. This contributes [SITUATION nonstative] to the definition of the complement of the progressive auxiliary. In the §5.6, I illustrate this in some detail for the English auxiliary (*be*).

5.5.3 Inflection

The infrastructure that handles inflectional morphology and creates lexical rules in the Matrix customization system was developed by O'Hara (2008). In all important respects it is not part of my work; however, it is integral to the implementation of tense and aspect. I describe it here in brief, general terms and give an example of its use handling tense and aspect related inflection.

In the Lexicon section of the questionnaire, inflectional morphemes can be defined for verbs, nouns and determiners. These morphemes are defined as occurring in optional or obligatory prefix or suffix slots. Interaction between slots can also be defined, i.e., the occurrence of a morpheme in one slot can be defined to place conditions on the occurrence of morphemes in another. The output of this inflectional infrastructure is lexical rules. See O'Hara (2008) for more detail.

Inflectional morphology definition interacts with tense and aspect in basically two ways. First, control over ordering and co-occurrence of morphemes is required to restrict strings to grammatical forms. For example, tense or aspect morphemes generally have one grammatical ordering; more than one tense morpheme on a verb may be ungrammatical; two morphemes might both be required to convey a particular tense or aspect.[66]

The second way that tense and aspect are associated with inflectional morphology in the questionnaire is through the definition of lexical rules. On the questionnaire, this translates to associating features with morphemes as they are defined. As an example, consider the fusional inflection of the French *imparfait* form illustrated in example (13b). The *-ait* inflec-

[64] In this case, assigning features/values directly to the auxiliary would be the same as constraining those values on its complement. This option is currently disallowed on the questionnaire strictly because it is redundant and does not seem as transparent to me.

[65] In addition, there is one more question in this section of the questionnaire about the subject of V or VP auxiliary complements but it is unrelated to the issues of tense and aspect.

[66] See O'Hara (2008) for her handling of the elaborate ordering and co-occurrence restrictions in Zulu.

tional suffix conveys third person singular number, past tense, and imperfective viewpoint aspect. Figure 14 is a screenshot of the part of the Inflection section of the Lexicon page where features are associated with that morpheme. This morpheme, among others, is associated with a particular morphological slot. Figure 15 contains the TDL for the corresponding lexical rule generated in the customized grammar.

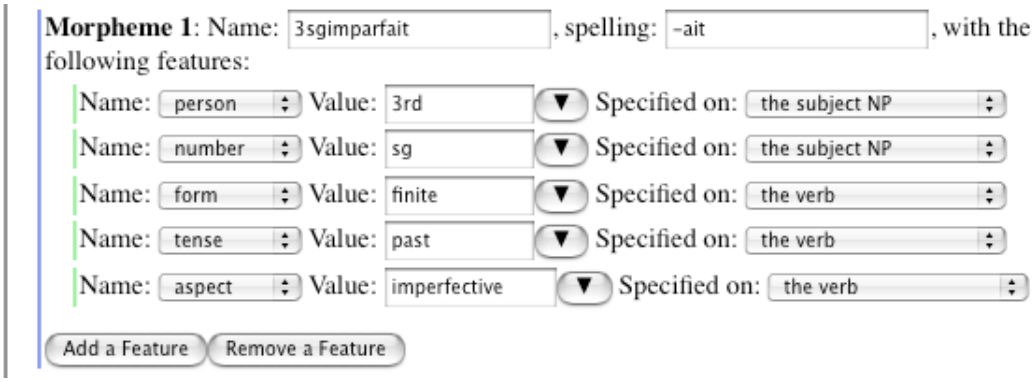


Figure 14: Questionnaire: associating features with a fusional morpheme

```
3sg-imparfait-lex-rule := main-verb-inflection-lex-rule &
      [SYNSEM.LOCAL[CAT[VAL.SUBJ.FIRST.LOCAL.CONT.HOOK.INDEX.PNG[NUM sg
                                                                 PER 3rd]
                        HEAD.FORM  finite]
                  CONT.HOOK.INDEX.E[TENSE  past
                                    ASPECT imperfective]]].
```

Figure 15: TDL: 3sg-imparfait (*-ait*) lexical rule

5.6 Extended Example

In this section, I present a single example in more detail. Specifically, I illustrate the definition of a model of the English Progressive construction. I provide details of the questionnaire choices, examples of TDL generated in the grammar the system creates based on the questionnaire choices and an MRS produced for a sentence parsed by the grammar in the LKB. What is explicated here would be part of a definition of the complete tense and aspect system which, in turn, would be part of a definition of an English grammar. The English Progressive is of interest for several reasons. The English Progressive construction can be used to demonstrate this implementation's handling of tense, aspect, auxiliaries, agreement with a subject noun phrase, inflection on both the auxiliary and the main verb, constraints on the form of the auxiliary's complement and constraints on the value of a semantic feature of the complement.

In the following, I walk through the choices made on the various relevant pages of the questionnaire to define the customized grammar. Note that there were many other choices made on the questionnaire in order create a grammar sufficient to parse strings like these, e.g., those choices necessary to define nouns and agreement feature values, which are not included in this discussion. In §5.6.2, I provide some sample results.

5.6.1 Questionnaire

On the Word Order page there are three choices to make regarding auxiliaries. English has auxiliaries and they occur before a VP complement. Figure 16 is a screenshot of the questionnaire filled out with these choices.

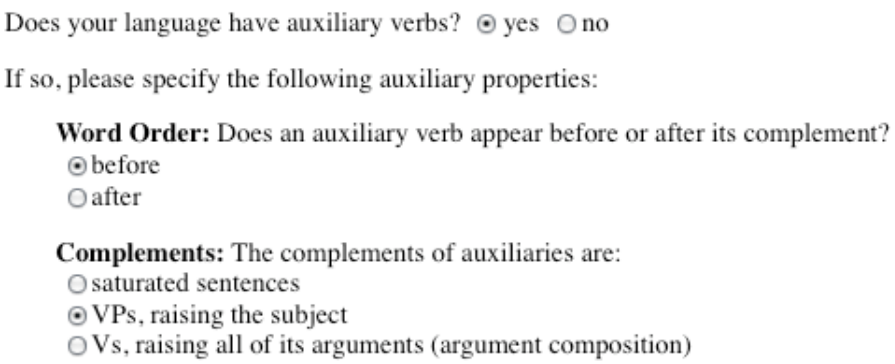


Figure 16: Extended example: Word Order choices

On the Tense and Aspect page, feature values are defined in hierarchies. The model of the English Progressive assumed in this example is basically consistent with Comrie (1976, 1985) although I make no claim that this particular analysis is definitive. Specifically, I assume that English has a semantic perfective/imperfective contrast although most sentences are underspecified for viewpoint aspect, that progressive is a marked subtype of imperfective, and that the primary tense contrast is past/non-past. In addition, I assume that stativity is a lexical property of some English verbs. This model is defined on the questionnaire through feature value hierarchies on the Tense and Aspect Features page.

The past/non-past tense contrast is defined through the choice of the values *past* and *nonpast* from the selection of common TENSE hierarchy elements provided. The viewpoint hierarchy is defined as values of the feature ASPECT. Figure 17 displays a screenshot of this section of the questionnaire.[67] The ASPECT values *pfv* (perfective) and *ipfv* (imperfective) inherit from the value *aspect* and *prog* (progressive) inherits from *ipfv*.

[67]Note that, while this hierarchy produces a grammar that correctly distinguishes grammatical and ungrammatical sentences, it overgenerates, i.e., will produce too many sentences when generating from semantic representations. This can be corrected several ways including through an expansion of this hierarchy. However, such an expansion would complicate multiple parts of this explanation without contributing anything useful to this discussion.

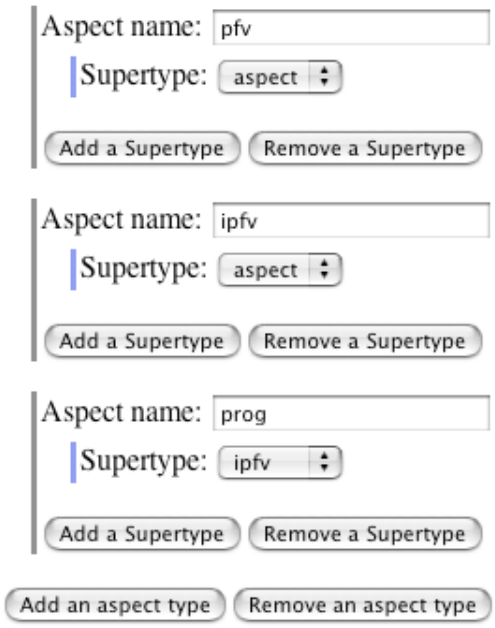


Figure 17: Extended example: Viewpoint hierarchy choices

The only situation aspect features defined for this example grammar are related to stativity. The SITUATION values *stat* (stative) and *nonstat* (non-stative) are defined as subtypes of the value *situation*. The final feature for which to define values on this page is the feature FORM. On the questionnaire, three subtypes of the FORM value *nonfinite* are defined: *infinitive*, *pstpart* (past participle) and *prspart* (present participle). Three subtypes are defined for illustrative purposes; only *prspart* is actually used by this example grammar. Figure 11 illustrates the definition of non-finite forms through the questionnaire. The only difference between the values defined on Figure 11 and the ones defined for this example grammar are the names of the values.

The Lexicon page contains all of the remaining tense and aspect related sections of the questionnaire. Verb types, the auxiliary and various inflectional morphemes are defined on this page. Four verb types are specifically defined for this example reflecting non-stative intransitive, non-stative transitive, stative transitive, and stative intransitive classes of verbs. Figure 18 is a screenshot of the questionnaire definition of one of these types: iv-nonstate (non-stative intransitive). This type is defined to have a *nonstat* value of the feature SITUATION, the argument structure is defined to be intransitive and a stem spelled *sleep* is defined. This stem will appear in the lexicon in the customized grammar inheriting from the supertype iv-nonstate.

Auxiliary verb types are defined separately from other verbs types as the properties that must be specified are different. A single auxiliary type is defined for this example grammar. Figure 19 contains the screenshot of the choices that comprise this definition.[68] This type

[68]The choices of subject type is not discussed here as it is not directly related to tense or aspect.

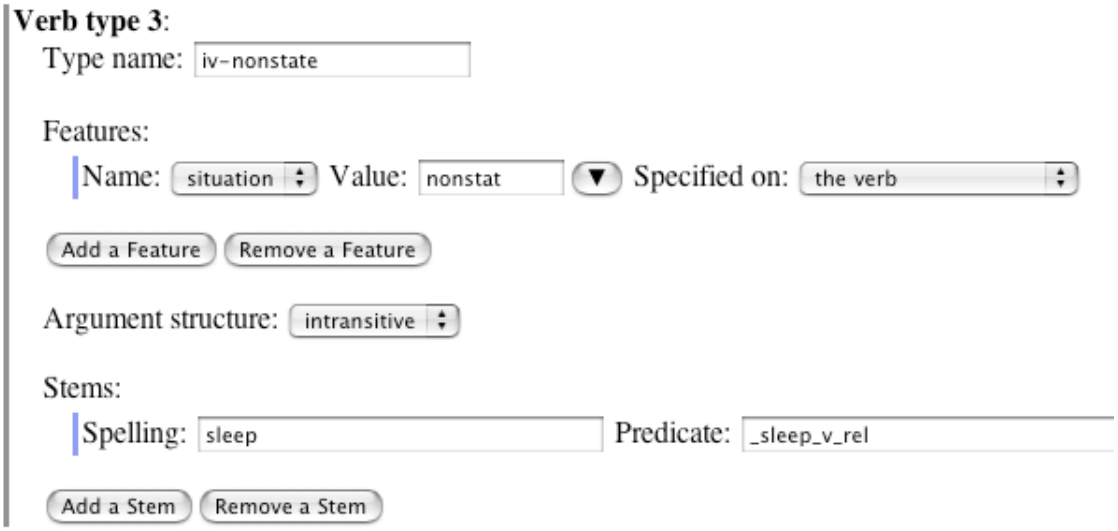


Figure 18: Questionnaire: non-stative intransitive verb type definition

contributes nothing to the semantics of a phrase other than tense and aspect so it is defined as contributing no predicate of its own. Constraining the complement of an auxiliary type specifies values of the verbal predicate. In this specific case, the FORM of the complement is constrained to be *prspart*. In conjunction with the definition of the -ing morpheme in the Verbal Inflection section, this will require the complement of a be-type auxiliary to be a present participle. In addition, the complement is specified to be progressive (ASPECT *prog*) and non-stative (SITUATION *nonstat*). The only stem defined for this auxiliary type is the auxiliary verb *be*. Figure 20 contains the TDL created in the customized example grammar from these choices.

The last step is the definition of inflectional morphology in the Verbal Inflection section. In order to parse an informative collection of sentences, multiple verb inflection morphemes are defined: 3rd person singular non-past (*-s*), non-3rd person non-past (∅), plural non-past (∅), past (*-ed*), and present participle (*-ing*). Each of these is a possible morpheme in an obligatory morpheme slot that attaches after any verb. Each morpheme definition includes its spelling, an obligatory FORM feature value and any other feature values associated with the morpheme. Figure 21 illustrates the choices made to define the 3rd person singular non-past inflectional morpheme *-s*. The *-ing* inflectional morpheme is simply associated with the *prspart* value of the FORM feature. This, in conjunction with the auxiliary definition, constrains complements of the auxiliary verb *be* to *-ing* forms.

All of the choices made on the questionnaire are retained by the customization system in a 'choices' file. This choices file is used by the system to create a customized grammar. The choices file for this extended example, designed to parse English Progressive sentences, can be found in Appendix B. The customized grammar produced consists of a collection of files. For example, there is a separate lexicon file for the defined stems. However, the majority of the language-specific code generated by the customization system is contained in one TDL

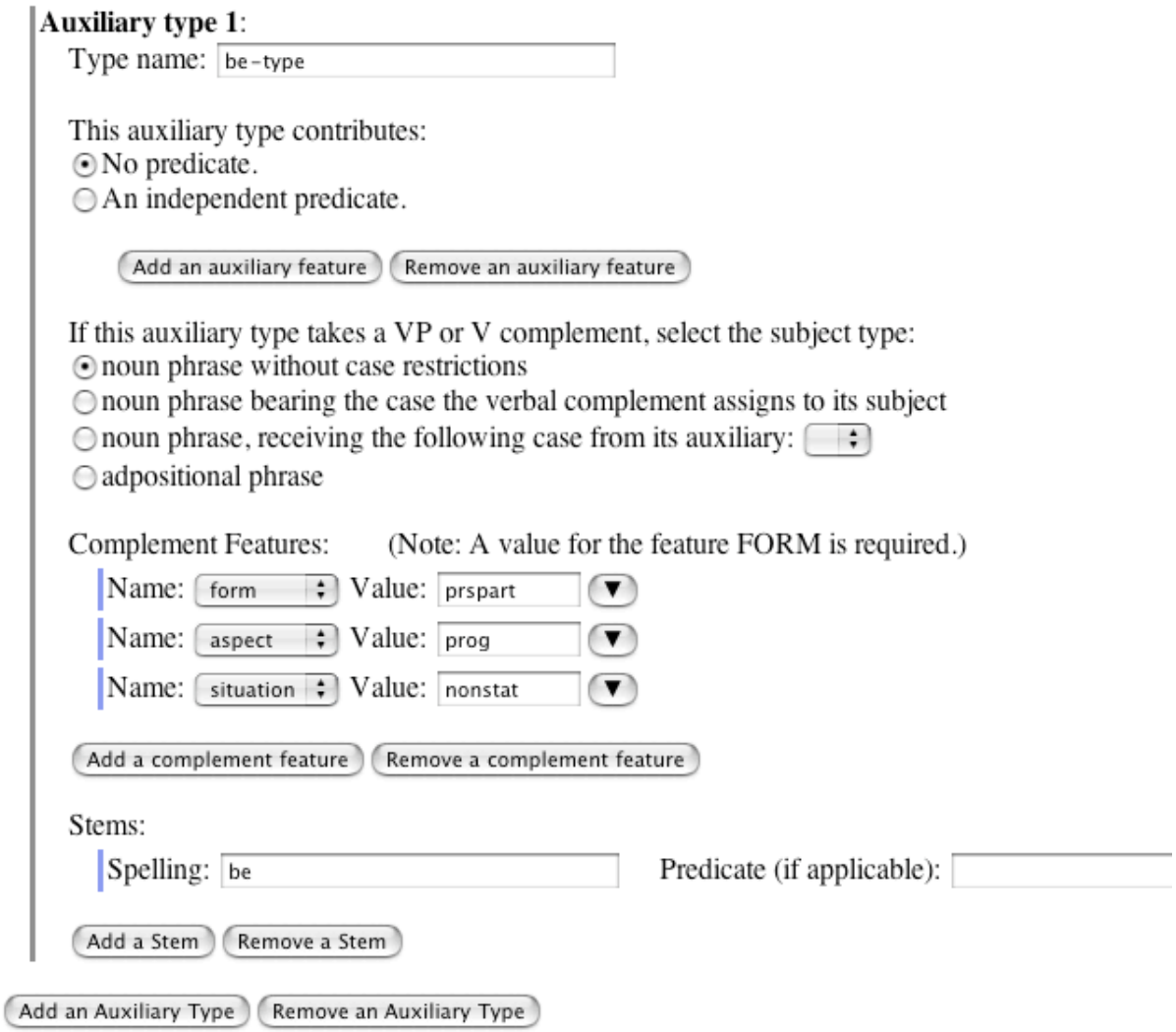


Figure 19: Extended example: auxiliary type choices

```
be-type-aux-lex := subj-raise-aux-no-pred &
  [SYNSEM.LOCAL.CAT.VAL.COMPS.FIRST.LOCAL[CAT.HEAD.FORM prspart,
                                          CONT.HOOK.INDEX.E[ASPECT prog,
                                                            SITUATION nonstat]]].
```

Figure 20: TDL: generated from auxiliary verb type choices

file.[69] Appendix C contains this file for the extended example discussed here and defined by the choices file in Appendix B.

[69]This file is named by the customization system after the language specified on the first page of the questionnaire.

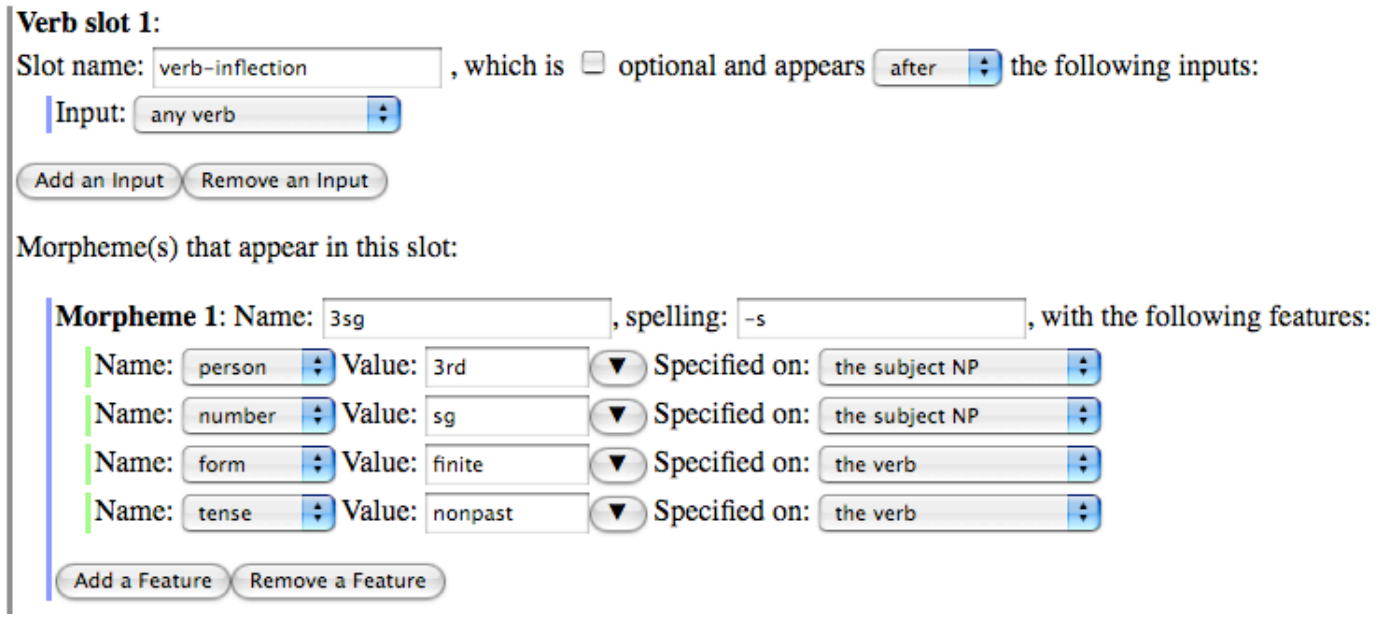


Figure 21: Extended example: morpheme definition choices

5.6.2 Sample results

The customized grammar created for this example is fairly minimal, focusing on the elements of the grammar needed to parse simple sentences illustrating the Progressive and demonstrating the various elements of the meta-modeling approach to tense and aspect assumed in this work. This small grammar is designed specifically to parse sentence (regularized strings of morphemes) such as `the cat be-s sleep-ing`, i.e., *The cat is sleeping*, while rejecting strings like *`the cat be-s know-ing the answer`. This small grammar also parses simple non-auxiliary constructions used for contrast in testing. Table 4 contains a selection of grammatical and ungrammatical test strings and, for each, the properties intended for the string.[70] The example grammar is fully accurate on these (and other) sentences, parsing the grammatical and not the ungrammatical and contributing the appropriate semantic values to the MRS.

The MRS in Figure 22 is a screenshot of the MRS display provided by the LKB parsing environment for the sentence `the cat be-s sleep-ing` parsed with the example grammar. This MRS displays the expected feature values for TENSE, ASPECT and SITUATION: *nonpast*, *prog* and *nonstat*, respectively.

5.7 Further Work

This implementation is designed to provide basic functionality and a foundation for further refinement — both further refinement by the users of starter-grammars created from the

[70] Properties in italics indicate the source of the ungrammaticality.

`the cat be-s sleep-ing`
intransitive; present; progressive ; nonstative; singular
`the cat be-ed sleep-ing`
intransitive; past; progressive; nonstative; singular
`the cat sleep-s`
intransitive; present; nonstative; singular
`the cat-s be chase-ing the dog`
transitive; present; progressive; nonstative; plural
∗`the cat be-s believe-ing the dog`
transitive; present; *progressive* ; *stative*; singular
∗`the cat be-s chase-ing`
transitive; present; progressive; nonstative; singular
∗`the cat sleep-ing`
intransitive; *present*; progressive ; nonstative; singular
∗`the cat be sleep-ing`
intransitive; *present*; progressive ; nonstative; singular
∗`the cats be-s sleep-ing`
intransitive; present; progressive ; nonstative; *plural*

Table 4: Extended example: sample test suite items

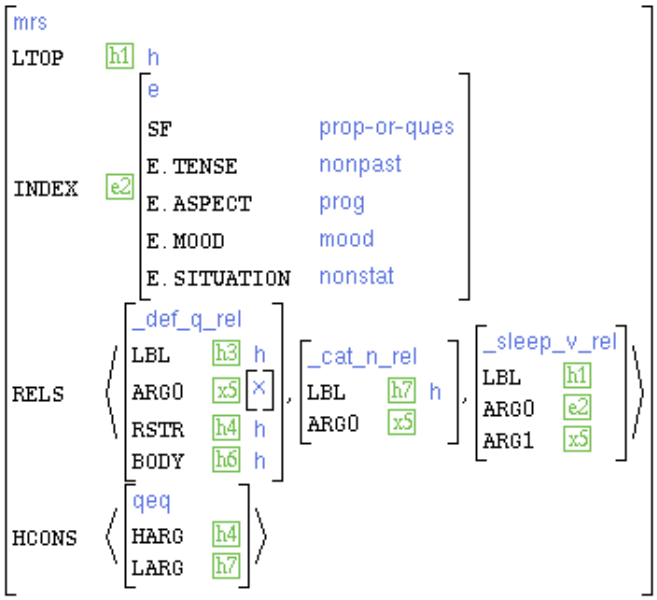


Figure 22: MRS: `the cat be-s sleep-ing`

customization and further refinement of the implementation itself in future Grammar Matrix development cycles. There are many avenues for continued development. Of these, there are essentially two types: questionnaire enhancements and phenomena expansion. In terms of

the questionnaire, there are several areas that can be improved. The most urgent need is for improvement to the process for defining lexical types and type hierarchies. The current implementation does not provide the user sufficient flexibility when defining type hierarchies resulting in a significant amount of duplicated effort when creating types. Questionnaire enhancement in this area is already underway. In addition, the Other Features page needs to be enhanced to allow a broader range of feature paths, thereby allowing other types of features to be defined.

It might also be useful to add some pre-defined ASPECT feature value hierarchies to the questionnaire. It might be worth providing an option for a pre-defined perfective/imperfective viewpoint contrast.[71] Likewise, it might be useful for some linguists to have the option to choose a pre-defined hierarchy of the situation aspect features: dynamic, durative and telic. This option may apply to a fairly limited number of languages or analyses but the effort to build a cross-classified hierarchy of these three values is not insignificant. Finally, some users might prefer binary features—while binary features are less useful for building analyses based on underspecification, they may have other advantages for a particular user.[72]

As this is a basic, preliminary implementation, there are limitations on the coverage of the system. Some of these limitations reflect the need to reduce the complexity of the problem in the initial implementation. Issues of nominal tense have been excluded for this reason. Likewise, any specific implementation of the semantics of the perfect has been ignored. In addition, aspectually complex structures like inceptives and terminatives, e.g., English *start to V* or *finish V-ing*, have also been excluded. The decision to address only tense and aspect instead of TAM as a whole, excluding mood and modality, was made, again, in an attempt to keep the scope manageable. These all represent areas for future work. Specifically, this implementation could easily be expanded to include a MOOD feature, expanding the coverage to accommodate the semantics of modals. A treatment of the semantics of the perfect could also be added through the introduction of another TAM feature (PERFECT), as in the implementation of the English Perfect (Dan Flickinger, p.c.) in the English Resource Grammar (ERG) (Flickinger, 2000). Although, it remains to be seen if a more elaborate implementation of the semantics of the perfect is cross-linguistically required. In addition, some treatment of evidentials, whether as a subset of mood or as a separate phenomenon, seems warranted, given languages like Turkish. In Turkish, the primary contrast is between direct experience and indirect or inferential evidence, not tense or aspect.

Another group of phenomena is excluded from this implementation due to a lack of support for them within the current customization system. The development of the Grammar Matrix customization system is, by design, incremental. Grammar engineering involves cycles of development, at each step ensuring compatibility with previous steps and progress (i.e. extension of coverage) before moving on. This is as true for the development of the Grammar Matrix as it is for individual grammars. The incremental development of the Matrix customization system began with core matrix clauses, and has not yet expanded into subor-

[71] This has not yet been provided since the effort it might save the user has seemed minimal.

[72] Users can define their own boolean values, but the current customization system does not expose the *boolean* type and its subtypes that are already part of the Matrix core grammar.

dinate clauses, adverbs, or serial verb constructions. This means that this implementation of tense and aspect does not address issues involved in subordinate, relative or complement clauses and in particular provides no support for relative tense or sequence of tense analyses. The fact that the customization system has no implementation of serial verb constructions provided more impetus for excluding inceptives and terminatives. In addition, the lack of support of adverbs has led to the exclusion of a significant class of tense and aspect particles that act like adverbial modifiers, e.g., English *up* as in *eat his lunch up* (See Jackendoff, 2002) from this implementation. However, development of tense and aspect, as well as other affected libraries, will necessarily continue as coverage of the broader customization system expands to incorporate these and other phenomena that involve interactions.

6 Conclusion

The implementation of tense and aspect in the Matrix customization system discussed here creates a platform for defining basic tense and aspect marking. The meta-modeling approach provides the user the flexibility to define language-specific features, types and hierarchies as well as to determine what contributes, and constrains, feature values. Specifically, this implementation supports the definition of a variety of tense and aspect system elements including:

- semantic features
- feature hierarchies
- lexical types with associated features
- lexical items with associated features
- auxiliaries with associated features
- feature-based constraints on auxiliary complements
- inflectional morphemes with associated features
- feature-based constraints on inflectional attachment

This implementation allows the user to handle common forms of tense and aspect morphology, to provide a reflection of them in the semantics and to use them to constrain ungrammaticality. It addresses an array of the most prominent phenomena associated with tense and aspect. While it represents only a first step in the development of tense and aspect in the Matrix customization system, it provides the user a significant jump-start towards the development of a tense and aspect system reflecting language-specific facts and linguist-specific analyses. Importantly, it also provides a platform for developing analyses of tense and aspect through experimentation with various hypotheses.

Acknowledgments

I would like to thank Emily Bender, Ann Copestake, Scott Drellishak, Dan Flickinger, Sharon Hargus and Fei Xia for helpful discussion. This material is based upon work supported by the National Science Foundation under Grant No. 0644097. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Anderson, G. D. S. (2006). *Auxiliary Verb Constructions*. Oxford Studies in Typology and Linguistic Theory. New York: Oxford University Press.

Axelrod, M. (1993). *The Semantics of Time: Aspectual Categorization in Koyukon Athabaskan*. Lincoln: University of Nebraska Press in cooperation with the American Indian Studies Research Institute, Indiana University, Bloomington.

Bach, E. W. (1981). On time, tense and aspect: an essay in english metaphysics. In P. Cole (Ed.) *Radical Pragmatics*, (pp. 62–81). Academic Press.

Bender, E. M. (2008). Grammar Engineering for Linguistic Hypothesis Testing. In N. Gaylord, A. Palmer, & E. Ponvert (Eds.) *Proceedings of the Texas Linguistics Society X Conference: Computational Linguistics for Less-Studied Languages*, (pp. 16–36). Stanford: CSLI Publications ONLINE.
URL `http://cslipublications.stanford.edu/TLS/TLS10-2006/TLS10_Bender.pdf`

Bender, E. M., Drellishak, S., Fokkens, A., Poulson, L., & Saleem, S. (2010). Grammar customization. *Research on Language & Computation*, (pp. 1–50). 10.1007/s11168-010-9070-1.
URL `http://dx.doi.org/10.1007/s11168-010-9070-1`

Bender, E. M., & Flickinger, D. (2005). Rapid prototyping of scalable grammars: Towards modularity in extensions to a language-independent core. In *Proceedings of the 2nd International Joint Conference on Natural Language Processing IJCNLP-05 (Posters/Demos)*. Jeju Island, Korea.

Bender, E. M., Flickinger, D., & Öepen, S. (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, & R. Sutcliffe (Eds.) *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, (pp. 8–14). Taipei, Taiwan.

Bender, E. M., Flickinger, D., & Oepen, S. (2011). Grammar engineering and linguistic hypothesis testing: Computational support for complexity in syntactic analysis. In E. M.

Bender, & J. E. Arnold (Eds.) *Language from a Cognitive Perspective: Grammar, Usage and Processing*, (pp. 5–29). Stanford, CA: CSLI Publications.

Bertinetto, P. M. (2000). The progressive in Romance, as compared with english. In Ö. Dahl (Ed.) *Tense and aspect in the languages of Europe*, (pp. 559–604). Berlin; New York: Mouton de Gruyter.

Bertinetto, P. M., & Delfitto, D. (2000). Aspect vs. Actionality: Why they should be kept apart. In Ö. Dahl (Ed.) *Tense and aspect in the languages of Europe*, (pp. 189–225). Berlin; New York: Mouton de Gruyter.

Binnick, R. I. (1991). *Time and the verb : a guide to tense and aspect*. New York: Oxford University Press.

Binnick, R. I. (2006). The Project on Annotated Bibliography of Contemporary Research in Tense, Grammatical Aspect, Aktionsart, and Related Areas.
URL `http://www.utsc.utoronto.ca/~binnick/TENSE/index.html`

Bybee, J. L., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: The University of Chicago Press.

Comrie, B. (1976). *Aspect: an introduction to the study of verbal aspect and related problems*. Cambridge; New York: Cambridge University Press.

Comrie, B. (1978). Ergativity. In W. P. Lehmann (Ed.) *Syntactic Typology: Studies in the Phenomenology of Language*, chap. 7, (pp. 329–394). Austin: University of Texas Press.

Comrie, B. (1985). *Tense*. Cambridge [Cambridgeshire]; New York: Cambridge University Press.

Copestake, A. (2001). *Implementing Typed Feature Structure Grammars*. Stanford: CSLI Publications.

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

Dahl, Ö. (1985). *Tense and Aspect Systems*. Oxford; New York: B. Blackwell.

Dahl, Ö. (2000). The tense-aspect systems of European languages in a typological perspective. In Ö. Dahl (Ed.) *Tense and aspect in the languages of Europe*, (pp. 3–26). Berlin; New York: Mouton de Gruyter.

Dahl, Ö., & Velupillai, V. (2008a). The Future Tense. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.) *The World Atlas of Language Structures Online*, chap. 67. Munich: Max Planck Digital Library.
URL `http://wals.info/feature/67`

Dahl, Ö., & Velupillai, V. (2008b). The Past Tense. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.) *The World Atlas of Language Structures Online*, chap. 66. Munich: Max Planck Digital Library.
URL `http://wals.info/feature/66`

Dahl, Ö., & Velupillai, V. (2008c). Perfective/Imperfective Aspect. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.) *The World Atlas of Language Structures Online*, chap. 65. Munich: Max Planck Digital Library.
URL `http://wals.info/feature/65`

de Haan, F. (1999). Evidentiality and Epistemic Modality: Setting Boundaries. *Southwest Journal of Linguistics*, *18*, 83–101.

de Swart, H. (1998). Aspect Shift and Coercion. *Natural Language & Linguistic Theory*, *16*(2), 347–385.

Dowty, D. R. (1979). *Word meaning and Montague grammar : the semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht; Boston: D. Reidel Pub. Co.

Drellishak, S. (2009). *Widespread but Not Universal: Improving the Typological Coverage of the Grammar Matrix*. Ph.D. thesis, University of Washington.

Drellishak, S., & Bender, E. M. (2005). A coordination module for a crosslinguistic grammar resource. In S. Müller (Ed.) *The Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, (pp. 108–128). Stanford: CSLI Publications.

Dryer, M. S. (2008). Position of Tense-Aspect Affixes. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.) *The World Atlas of Language Structures Online*, chap. 69. Munich: Max Planck Digital Library.
URL `http://wals.info/feature/69`

Ebert, K. H. (2000). Aspect in maltese. In Ö. Dahl (Ed.) *Tense and aspect in the languages of Europe*, (pp. 753–785). Berlin: Mouton de Gruyter.

Faust, N. (1978). *Gramática cocama: Lecciones para el aprendizaje del idioma cocamama: Lecciones para el aprendizaje del idioma cocama*, vol. 6 of *Lingüística Peruana*. Yarinacocha: Instituto Lingüístico de Verano, 2 ed.

Flickinger, D. (2000). On Building a More Efficient Grammar by Exploiting Types. *Natural Language Engineering*, *6 (1) (Special Issue on Efficient Processing with HPSG)*, 15 – 28.

Fokkens, A. (2010). Documentation for the Grammar Matrix word order library. Tech. rep., Saarland University.

Frawley, W. (1992). *Linguistic semantics*. Hillsdale, N.J.: L. Erlbaum Associates.

Givón, T. (1984). *Syntax: a functional-typological introduction*. Amsterdam; Philadelphia: John Benjamins.

Golla, V. (2007). Linguistic prehistory. In T. L. Jones, & K. Klar (Eds.) *California prehistory: Colonization, Culture, and Complexity*, chap. 6, (pp. 71–82). Lanham, MD: AltaMira Press.

Haspelmath, M. (2008). Comparative concepts and descriptive categories in cross-linguistic studies. Unpublished manuscript.

Hopper, P. J. (1982). Introduction. In P. J. Hopper (Ed.) *Tense-aspect: between semantics & pragmatics*. Amsterdam; Philadelphia: John Benjamins.

Hoxie, F. E. (1996). *Encyclopedia of North American Indians*. Boston: Houghton Mifflin Company.

Hyman, L. M. (1980). Relative time reference in the bamileke tense system. *Studies in African Linguistics*, *11*, 227–237.

Hymes, D. (1975). From space to time in tenses in kiksht. *International Journal of American Linguistics*, *41*, 313–329.

Jackendoff, R. (2002). English particle constructions, the lexicon, and the autonomy of syntax. In N. Dehé, R. Jackendoff, A. McIntyre, & S. Urban (Eds.) *Verb-particle explorations*, (pp. 67–94). Berlin; New York: Mouton de Gruyter.

Jakobson, R., & Waugh, L. R. (1984). *Russian and Slavic grammar: studies 1931 - 1981*. Berlin: Mouton.

Johanson, L. (2000). Viewpoint operators in European languages. In Ö. Dahl (Ed.) *Tense and aspect in the languages of Europe*, (pp. 27–188). Berlin; New York: Mouton de Gruyter.

Kibort, A. (2008). Grammatical features: Organisation of the feature inventory. Accessed 12-September-2009.
URL `http://www.grammaticalfeatures.net/inventory.html`

Kiparsky, P. (1998). Partitive case and aspect. In M. Butt, & W. Geuder (Eds.) *The projection of arguments: Lexical and compositional factors*. Stanford, CA: CSLI Publications.

Klein, W. (1994). *Time in language*. London; New York: Routledge.

Koenig, J.-P., & Muansuwan, N. (2000). How to End Without Ever Finishing: Thai Semiperfectivity. *Journal of Semantics*, *17*, 147–184.

Krieger, H.-U., & Schäfer, U. (1994). TDL - A Type Description Language for constraint-based grammarsanguage for constraint-based grammars. In *Proceedings of the 15th International Conference on Computational Linguistics*, (pp. 893–899). Kyoto, Japan.

Krifka, M. (1989). Nominal Reference, Temporal Constitution and Quantification in Event Semantics. In R. Bartsch, J. van Benthem, & P. von Emde Boas (Eds.) *Semantics and Contextual Expression*. Dordrecht: Foris Publication.

Lewis, M. P. (Ed.) (2009). *Ethnologue: Languages of the World, Sixteenth edition*. Dallas, TX: SIL International. Online version.
URL `http://www.ethnologue.com`

Lindstedt, J. (2001). Tense and Aspect. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.) *Language Typology and Language Universals*, (pp. 768–783). Berlin: Walter de Gruyter.

Malotki, E. (1983). *Hopi time: a linguistic analysis of the temporal concepts in the Hopi language*. Berlin; New York: Mouton Publishers.

Midgette, S. (1987). *The Navajo Progressive in Discourse Context: A Study in Temporal Semantics*. Ph.D. thesis, University of New Mexico.

Muansuwan, N. (2002). *Verb Complexes in Thai*. Ph.D. thesis, University at Buffalo, The State University of New York.

Mughazy, M. (2005). Rethinking Lexcial Aspect in Egyptian Arabic. In M. T. Alhawary, & E. Benmamoun (Eds.) *Perspectives on Arabic linguistics XVII-XVIII: papers from the seventeenth and eighteenth annual symposia on Arabic linguistics*, (pp. 133–172). Amsterdam; Philadelphia: J. Benjamins Pub. Co.

Nordlinger, R., & Sadler, L. (2004). Nominal tense in a crosslinguistic perspective. *Language*, *80*(4), 776–806.

Ogihara, T. (1998). Tense, aspect, and argument structure. In *Proceedings from Semantics and Linguistic Theory VIII (SALT8)*, (pp. pp. 169–184). Cornell University: CLC Publications.

O'Hara, K. (2008). *A Morphosyntactic Infrastructure for a Grammar Customization System*. Master's thesis, University of Washington, Seattle.

Paducheva, E., & Pentus, M. (2008). Formal and informal semantics of telicity. In S. D. Rothstein (Ed.) *Theoretical and crosslinguistic approaches to the semantics of aspect*, chap. 6, (pp. 191–216). Amsterdam; Philadelphia: John Benjamins Publishing.

Palmer, F. R. (1986). *Mood and Modality*. Cambridge University Press.

Parsons, T. (1990). *Events in the Semantics of English: A study in subatomic semantics*. Cambridge, MA: MIT Press.

Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago, Illinois: University of Chicago Press and CSLI Publications.

Reichenbach, H. (1947). *Elements of Symbolic Logic*. New York: Free Press.

Rice, K. (1989). *A grammar of Slave*. Berlin: Mouton de Gruyter.

Rice, K. (2000). *Morpheme order and semantic scope: word formation in the Athapaskan verb*. Cambridge; New York: Cambridge University Press.

Rothstein, S. D. (2008). Telicity, atomicity and the vendler classification of verbs. In S. D. Rothstein (Ed.) *Theoretical and crosslinguistic approaches to the semantics of aspect*, (pp. 43–78). Amsterdam; Philadelphia: John Benjamins Publishing.

Sag, I. A., Wasow, T., & Bender, E. M. (2003). *Synactic Theory: A Formal Introduction*. Stanford, CA: CSLI, second ed.

Saleem, S., & Bender, E. M. (2010). Argument optionality in the LinGO Grammar Matrix. In *Coling 2010: Posters*, (pp. 1068–1076). Beijing, China: Coling 2010 Organizing Committee.
URL `http://www.aclweb.org/anthology/C10-2123`

Sasse, H.-J. (2002). Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology*, *6*(2), 199–271.

Slabakova, R. (2004). Effect of Perfective Prefixes on Object Interpretation: A Theoretical and Empirical Issue. *Cahiers linguistiques d'Ottowa*, *32*, 122–142.

Smith, C. S. (1991). *The parameter of aspect*. Dordrecht; Boston: Kluwer Academic Publishers, 1st ed.

Smith, C. S. (1997). *The parameter of aspect*. Dordrecht; Boston: Kluwer, 2nd ed.

Smith, C. S. (2005). Time with and without tense. Paper presented at the International Round Table on Tense and Modality, Paris.

Smith, C. S. (2007). Tense and temporal interpretation. *Lingua*, *117*(2), 419 – 436.

Thieroff, R. (2000). On the areal distribution of tense-aspect categories in Europe. In Ö. Dahl (Ed.) *Tense and Aspect in the Languages of Europe*, (pp. 265–305). Berlin: Mouton de Gruyter.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, *66*(2), 143–160. (Reprinted in a revised version in Vendler, 1967).

Vlach, F. (1981). The Semantics of the Progressive. In P. J. Tedeschi, & A. E. Zaenen (Eds.) *Tense and Aspect*, vol. 14 of *Syntax and Semantics*, (pp. 271–292). New York: Academic Press.

Wilhelm, A. (2007). *Telicity and Durativity: A Study of Aspect in Dëne Sųłiné (Chipewyan) and German*. New York: Routledge.

Zeitoun, E., & Huang, M.-c. (2000). Concerning ka-, an overlooked marker of verbal derivation in Formosan languages. *Oceanic Linguistics*, *39*(2), 391–414.

A Appendix: Gloss Abbreviations

The following gloss abbreviations are used in this document. They are based on the Leipzig Glossing Rules available at: `http://www.eva.mpg.de/lingua/resources/glossing-rules.php`

1sg	first person singular
3sg	third person singular
ABS	absolutive
ACC	accusative
AUX	auxiliary
DAT	dative
ERG	ergative
FEM	feminine
FUT	future
IPFV	imperfective
NOM	nominative
PART	partitive
PFV	perfective
PROG	progressive
PRS	present
PST	past

B Appendix: Extended Example Choices File

```
version=18

section=general
language=Extended Example
archive=no

section=word-order
word-order=svo
has-dets=yes
noun-det-order=det-noun
has-aux=yes
aux-comp-order=before
aux-comp=vp
```

```
section=number
  number1_name=sg
  number2_name=pl

section=person
person=1-2-3
first-person=none

section=gender

section=case
case-marking=none

section=direct-inverse

section=tense-aspect
tense-definition=choose
past=on
nonpast=on
  aspect1_name=pfv
    aspect1_supertype1_name=aspect
  aspect2_name=ipfv
    aspect2_supertype1_name=aspect
  aspect3_name=prog
    aspect3_supertype1_name=ipfv
  situation1_name=stat
    situation1_supertype1_name=situation
  situation2_name=nonstat
    situation2_supertype1_name=situation
  nf-subform1_name=prspart
  nf-subform2_name=infinitive
  nf-subform3_name=pstpart

section=other-features

section=sentential-negation

section=coordination

section=matrix-yes-no

section=arg-opt
```

```
section=lexicon
  noun1_name=common
    noun1_feat1_name=person
    noun1_feat1_value=3rd
  noun1_det=obl
    noun1_stem1_orth=cat
    noun1_stem1_pred=_cat_n_rel
    noun1_stem2_orth=dog
    noun1_stem2_pred=_dog_n_rel
  noun-slot1_name=num
  noun-slot1_order=after
    noun-slot1_input1_type=noun1
    noun-slot1_morph1_name=singular
      noun-slot1_morph1_feat1_name=number
      noun-slot1_morph1_feat1_value=sg
    noun-slot1_morph2_name=plural
    noun-slot1_morph2_orth=-s
      noun-slot1_morph2_feat1_name=number
      noun-slot1_morph2_feat1_value=pl
  verb1_name=iv
  verb1_valence=intrans
  verb2_name=tv
  verb2_valence=trans
  verb3_name=iv-nonstate
    verb3_feat1_name=situation
    verb3_feat1_value=nonstat
    verb3_feat1_head=verb
  verb3_valence=intrans
    verb3_stem1_orth=sleep
    verb3_stem1_pred=_sleep_v_rel
  verb4_name=tv-nonstate
    verb4_feat1_name=situation
    verb4_feat1_value=nonstat
    verb4_feat1_head=verb
  verb4_valence=trans
    verb4_stem1_orth=chase
    verb4_stem1_pred=_chase_v_rel
  verb5_name=tv-state
    verb5_feat1_name=situation
    verb5_feat1_value=stat
    verb5_feat1_head=verb
```

```
verb5_valence=trans
  verb5_stem1_orth=believe
  verb5_stem1_pred=_believe_v_rel
verb6_name=iv-state
  verb6_feat1_name=situation
  verb6_feat1_value=stat
  verb6_feat1_head=verb
verb6_valence=intrans
  verb6_stem1_orth=exist
  verb6_stem1_pred=_exist_v_rel
aux1_name=be-type
aux1_sem=no-pred
aux1_subj=np
  aux1_compfeature1_name=form
  aux1_compfeature1_value=prspart
  aux1_compfeature2_name=aspect
  aux1_compfeature2_value=prog
  aux1_compfeature3_name=situation
  aux1_compfeature3_value=nonstat
  aux1_stem1_orth=be
verb-slot1_name=inflection
verb-slot1_order=after
  verb-slot1_input1_type=verb
  verb-slot1_morph1_name=3sg
  verb-slot1_morph1_orth=-s
    verb-slot1_morph1_feat1_name=person
    verb-slot1_morph1_feat1_value=3rd
    verb-slot1_morph1_feat1_head=subj
    verb-slot1_morph1_feat2_name=number
    verb-slot1_morph1_feat2_value=sg
    verb-slot1_morph1_feat2_head=subj
    verb-slot1_morph1_feat3_name=form
    verb-slot1_morph1_feat3_value=finite
    verb-slot1_morph1_feat3_head=verb
    verb-slot1_morph1_feat4_name=tense
    verb-slot1_morph1_feat4_value=nonpast
    verb-slot1_morph1_feat4_head=verb
  verb-slot1_morph2_name=pl
    verb-slot1_morph2_feat1_name=number
    verb-slot1_morph2_feat1_value=pl
    verb-slot1_morph2_feat1_head=subj
    verb-slot1_morph2_feat2_name=form
```

```
      verb-slot1_morph2_feat2_value=finite
      verb-slot1_morph2_feat2_head=verb
      verb-slot1_morph2_feat3_name=tense
      verb-slot1_morph2_feat3_value=nonpast
      verb-slot1_morph2_feat3_head=verb
    verb-slot1_morph3_name=non-3rd
      verb-slot1_morph3_feat1_name=person
      verb-slot1_morph3_feat1_value=1st, 2nd
      verb-slot1_morph3_feat1_head=subj
      verb-slot1_morph3_feat2_name=form
      verb-slot1_morph3_feat2_value=finite
      verb-slot1_morph3_feat2_head=verb
      verb-slot1_morph3_feat3_name=tense
      verb-slot1_morph3_feat3_value=nonpast
      verb-slot1_morph3_feat3_head=verb
    verb-slot1_morph4_name=past
    verb-slot1_morph4_orth=-ed
      verb-slot1_morph4_feat1_name=form
      verb-slot1_morph4_feat1_value=finite
      verb-slot1_morph4_feat1_head=verb
      verb-slot1_morph4_feat2_name=tense
      verb-slot1_morph4_feat2_value=past
      verb-slot1_morph4_feat2_head=verb
    verb-slot1_morph5_name=present-part
    verb-slot1_morph5_orth=-ing
      verb-slot1_morph5_feat1_name=form
      verb-slot1_morph5_feat1_value=prspart
      verb-slot1_morph5_feat1_head=verb
  det1_name=definite
    det1_stem1_orth=the
    det1_stem1_pred=_def_q_rel

section=test-sentences
sentence1=the cat be-s sleep-ing
sentence2=the cats be chase-ing the dog
```

C Appendix: Extended Example TDL File

```
;;; Grammar of Extended Example
;;; created at:
;;;     Wed Sep 09 22:12:31 UTC 2009
;;; based on Matrix customization system version of:
;;;     Tue Jul  7 16:33:47 UTC 2009

num-lex-rule := lexeme-to-word-rule & add-only-no-ccont-rule &
  [ DTR common-noun-lex ].

common-noun-lex := noun-lex &
  [ INFLECTED -,
    SYNSEM.LOCAL.CONT.HOOK.INDEX.PNG.PER 3rd ].

singular-lex-rule := const-ltow-rule & num-lex-rule &
  [ SYNSEM.LOCAL.CONT.HOOK.INDEX.PNG.NUM sg ].

plural-lex-rule := infl-ltow-rule & num-lex-rule &
  [ SYNSEM.LOCAL.CONT.HOOK.INDEX.PNG.NUM pl ].

inflection-lex-rule := lexeme-to-word-rule & add-only-no-ccont-rule &
  [ DTR verb-lex ].

verb-lex := lex-item &
  [ INFLECTED -,
    SYNSEM.LOCAL.CAT.HEAD verb ].

3sg-lex-rule := infl-ltow-rule & inflection-lex-rule &
  [ SYNSEM.LOCAL [ CAT [ VAL.SUBJ.FIRST.LOCAL.CONT.HOOK.INDEX.PNG [ PER 3rd,
                                                                    NUM sg ],
                         HEAD.FORM finite ],
                   CONT.HOOK.INDEX.E.TENSE nonpast ] ].

pl-lex-rule := const-ltow-rule & inflection-lex-rule &
  [ SYNSEM.LOCAL [ CAT [ VAL.SUBJ.FIRST.LOCAL.CONT.HOOK.INDEX.PNG.NUM pl,
                         HEAD.FORM finite ],
                   CONT.HOOK.INDEX.E.TENSE nonpast ] ].

non-3rd-lex-rule := const-ltow-rule & inflection-lex-rule &
  [ SYNSEM.LOCAL [ CAT [ VAL.SUBJ.FIRST.LOCAL.CONT.HOOK.INDEX.PNG.PER non-3rd,
                         HEAD.FORM finite ],
```

```
                    CONT.HOOK.INDEX.E.TENSE nonpast ] ].

past-lex-rule := infl-ltow-rule & inflection-lex-rule &
  [ SYNSEM.LOCAL [ CAT.HEAD.FORM finite,
                   CONT.HOOK.INDEX.E.TENSE past ] ].

present-part-lex-rule := infl-ltow-rule & inflection-lex-rule &
  [ SYNSEM.LOCAL.CAT.HEAD.FORM prspart ].

;;; Lexical types

;;; Nouns

noun-lex := basic-noun-lex & basic-one-arg & no-hcons-lex-item &
  [ ARG-ST < #spr >,
    SYNSEM.LOCAL.CAT.VAL [ COMPS < >,
                           SUBJ < >,
                           SPEC < >,
                           SPR < #spr &
                                 [ LOCAL.CAT.HEAD det,
                                   OPT - ] > ] ].

;;; Verbs

head :+ [ AUX bool,
    FORM form ].

main-verb-lex := verb-lex & basic-verb-lex &
  [ SYNSEM.LOCAL [ CAT [ HEAD.AUX -,
                         VAL [ SPR < >,
                               SPEC < >,
                               SUBJ < #subj > ] ],
                   CONT.HOOK.XARG #xarg ],
    ARG-ST.FIRST #subj &
                 [ LOCAL [ CAT.VAL [ SPR < >,
                                     COMPS < > ],
                           CONT.HOOK.INDEX #xarg ] ] ].

aux-lex := verb-lex &
  [ SYNSEM.LOCAL.CAT.HEAD.AUX + ].

intransitive-verb-lex := main-verb-lex & intransitive-lex-item &
```

```
  [ SYNSEM.LOCAL.CAT.VAL.COMPS < >,
    ARG-ST.FIRST.LOCAL.CAT.HEAD noun ].

transitive-verb-lex := main-verb-lex & transitive-lex-item &
  [ SYNSEM.LOCAL.CAT.VAL.COMPS < #comps >,
    ARG-ST < [ LOCAL.CAT.HEAD noun ],
             #comps &
             [ LOCAL.CAT [ VAL [ SPR < >,
                                 COMPS < > ],
                           HEAD noun ] ] > ].

iv-verb-lex := intransitive-verb-lex.

tv-verb-lex := transitive-verb-lex.

iv-nonstate-verb-lex := intransitive-verb-lex &
  [ SYNSEM.LOCAL.CONT.HOOK.INDEX.E.SITUATION nonstat ].

tv-nonstate-verb-lex := transitive-verb-lex &
  [ SYNSEM.LOCAL.CONT.HOOK.INDEX.E.SITUATION nonstat ].

tv-state-verb-lex := transitive-verb-lex &
  [ SYNSEM.LOCAL.CONT.HOOK.INDEX.E.SITUATION stat ].

iv-state-verb-lex := intransitive-verb-lex &
  [ SYNSEM.LOCAL.CONT.HOOK.INDEX.E.SITUATION stat ].

;;; Auxiliaries

subj-raise-aux := aux-lex & trans-first-arg-raising-lex-item &
  [ SYNSEM.LOCAL [ CAT.VAL [ SUBJ < #subj >,
                             COMPS < #comps >,
                             SPR < >,
                             SPEC < > ],
                   CONT.HOOK.XARG #xarg ],
    ARG-ST < #subj &
             [ LOCAL [ CONT.HOOK.INDEX #xarg,
                       CAT [ VAL [ SPR < >,
                                   COMPS < > ],
                             HEAD noun ] ] ],
             #comps &
             [ LOCAL.CAT [ VAL [ SUBJ < [  ] >,
```

```
                                  COMPS < > ],
                               HEAD verb ] ] > ].

subj-raise-aux-no-pred := subj-raise-aux & raise-sem-lex-item &
  [ ARG-ST < [  ],
             [ LOCAL.CAT.HEAD.AUX - ] > ].

be-type-aux-lex := subj-raise-aux-no-pred &
  [ SYNSEM.LOCAL.CAT.VAL.COMPS.FIRST.LOCAL [ CAT.HEAD.FORM prspart,
                                             CONT.HOOK.INDEX.E [ ASPECT prog,
                                                                 SITUATION nonstat ] ] ].

;;; Determiners
;;; SPEC is non-empty, and already specified by basic-determiner-lex.

determiner-lex := basic-determiner-lex & basic-zero-arg &
  [ SYNSEM.LOCAL.CAT.VAL [ SPR < >,
                           COMPS < >,
                           SUBJ < > ] ].

definite-determiner-lex := determiner-lex.

png :+ [ PER person,
    NUM number ].

;;; Person

person := *top*.
3rd := person.
non-3rd := person.
2nd := non-3rd.
1st := non-3rd.

;;; Number

number := *top*.
sg := number.
pl := number.

;;; Form

form := *top*.
```

```
nonfinite := form.
finite := form.
prspart := nonfinite.
infinitive := nonfinite.
pstpart := nonfinite.

;;; Tense

nonpast := tense.
past := tense.

;;; Aspect

pfv := aspect.
ipfv := aspect.
prog := ipfv.

situation := sort.

tam :+ [ SITUATION situation ].

;;; Situation

stat := situation.
nonstat := situation.

;;; Phrasal types

basic-head-comp-phrase :+ [ SYNSEM.LOCAL.CAT.MC #mc,
    HEAD-DTR.SYNSEM.LOCAL.CAT.MC #mc ].

basic-head-mod-phrase-simple :+ [ SYNSEM.LOCAL.CAT.MC #mc,
    NON-HEAD-DTR.SYNSEM.LOCAL.CAT.MC #mc ].

head-comp-phrase := basic-head-1st-comp-phrase & head-initial.

subj-head-phrase := decl-head-subj-phrase & head-final &
  [ HEAD-DTR.SYNSEM.LOCAL.CAT.VAL.COMPS < > ].

; Rules for building NPs.  Note that the Matrix uses SPR for
; the specifier of nouns and SUBJ for the subject (specifier) of verbs.
```

```
head-spec-phrase := basic-head-spec-phrase & head-final.

; Bare NP phrase.  Consider modifying the PRED value of the quantifier relation
; introduced to match the semantic effect of bare NPs in your language.

bare-np-phrase := basic-bare-np-phrase &
  [ C-CONT.RELS <! [ PRED "exist_q_rel" ] !> ].
```