A Universal Feature Schema for Rich Morphological Annotation and Fine-Grained Cross-Lingual Part-of-Speech Tagging

John Sylak-Glassman[†], Christo Kirov[†], Matt Post[‡,§], Roger Que[§], and David Yarowsky[§⋆]

[†]Center for Language and Speech Processing
[‡] Human Language Technology Center of Excellence
[§]Department of Computer Science
Johns Hopkins University
Baltimore, MD 21218
jcsg@jhu.edu, ckirov@gmail.com, post@cs.jhu.edu, query@jhu.edu, yarowsky@jhu.edu

Abstract. Semantically detailed and typologically-informed morphological analysis that is broadly applicable cross-linguistically has the potential to improve many NLP applications, including machine translation, n-gram language models, information extraction, and co-reference resolution. In this paper, we present a universal morphological feature schema, which is a set of features that represent the finest distinctions in meaning that are expressed by inflectional morphology across languages. We first present the schema's guiding theoretical principles, construction methodology, and contents. We then present a method of measuring cross-linguistic variability in the semantic distinctions conveyed by inflectional morphology along the multiple dimensions spanned by the schema. This method relies on representing inflected wordforms from many languages in our universal feature space, and then testing for agreement across multiple aligned translations of pivot words in a parallel corpus (the Bible). The results of this method are used to assess the effectiveness of cross-linguistic projection of a multilingual consensus of these fine-grained morphological features, both within and across language families. We find high cross-linguistic agreement for a diverse range of semantic dimensions expressed by inflectional morphology.

Keywords: inflectional morphology · linguistic typology · universal schema · cross-linguistic projection

1 Introduction

Semantically detailed and typologically-informed morphological analysis that is broadly applicable cross-linguistically has the potential to improve many NLP applications, including machine translation (particularly of morphologically rich

⋆ The first two authors contributed equally to this paper.

languages), n-gram language models, information extraction (particularly event extraction), and co-reference resolution.

In this paper, we first present a novel universal morphological feature schema. This schema is a set of features that represent the finest distinctions in meaning that are expressed by inflectional morphology across languages. The purpose of the proposed universal morphological feature schema is to allow any given overt, affixal (non-root) inflectional morpheme in any language to be given a precise, language-independent, semantically accurate definition.

As a demonstration of the utility and consistency of our universal schema, we show how it can enable cross-linguistic projection-based approaches to detailed semantic tagging. We measure the cross-linguistic variability in the semantic distinctions conveyed by inflectional morphology along multiple dimensions captured by our schema. This method relies on representing inflected wordforms from many languages in our universal feature space, and then testing for feature agreement across multiple translations of pivot words chosen from a parallel text (e.g., the Bible). We find high cross-linguistic agreement for a diverse range of semantic dimensions expressed by inflectional morphology, both within and across language families. This is true even in some cases where we expect languages to diverge due to non-semantic or arbitrary divisions of the semantic space (e.g., when assigning grammatical gender to inanimate objects).

2 A Universal Morphological Feature Schema

This section describes the principles that inform the composition of the schema, the methodology used to construct it, and its contents. See Table 1 for a summary of the full schema that includes both the dimensions of meaning and their respective features.

2.1 Guiding Theoretical Principles

The purpose of the universal morphological feature schema is to allow any given overt, affixal (non-root) inflectional morpheme in any language to be given a precise, language-independent, semantically accurate definition. This influences the overall architecture of the schema in two significant ways.

First, the schema is responsible for capturing only the meanings of overt, non-root, affixal inflectional morphemes, which considerably limits the semantic-conceptual space that must be formally described using these features. This significant limitation of the range of data that must be modeled makes an interlingual approach to the construction of the schema feasible (as also noted by Sagot and Walther [47]).

Second, the schema is sensitive only to semantic content, not to overt surface form. This follows the insight in linguistic typology that "crosslinguistic comparison [...] cannot be based on formal patterns (because these are too diverse), but [must] be based primarily on universal conceptual-semantic concepts [27, p. 665, and references therein]. Due to the semantic focus of the schema, it contains no

features for indicating the form that a morpheme takes. Instead, the schema's features can be integrated into existing frameworks that can indicate the form of morphemes, such as Sagot and Walther [47] for NLP and the Leipzig Glossing Rules for theoretical and descriptive linguistics [12].

Dimension	*Features*
Aktionsart	ACCMP, ACH, ACTY, ATEL, DUR, DYN, PCT, SEMEL, STAT, TEL
Animacy	ANIM, HUM, INAN, NHUM
Aspect	HAB, IPFV, ITER, PFV, PRF, PROG, PROSP
Case	ABL, ABS, ACC, ALL, ANTE, APPRX, APUD, AT, AVR, BEN, CIRC, COM, COMPV, DAT, EQU, ERG, ESS, FRML, GEN, IN, INS, INTER, NOM, NOMS, ON, ONHR, ONVR, POST, PRIV, PROL, PROPR, PROX, PRP, PRT, REM, SUB, TERM, VERS, VOC
Comparison	AB, CMPR, EQT, RL, SPRL
Definiteness	DEF, INDEF, NSPEC, SPEC
Deixis	ABV, BEL, DIST, EVEN, MED, NVIS, PROX, REF1, REF2, REM, VIS
Evidentiality	ASSUM, AUD, DRCT, FH, HRSY, INFER, NFH, NVSEN, QUOT, RPRT, SEN
Finiteness	FIN, NFIN
Gender	BANTU1-23, FEM, MASC, NAKH1-8, NEUT
Information Structure	FOC, TOP
Interrogativity	DECL, INT
Mood	ADM, AUNPRP, AUPRP, COND, DEB, IMP, IND, INTEN, IRR, LKLY, OBLIG, OPT, PERM, POT, PURP, REAL, SBJV, SIM
Number	DU, GPAUC, GRPL, INVN, PAUC, PL, SG, TRI
Parts of Speech	ADJ, ADP, ADV, ART, AUX, CLF, COMP, CONJ, DET, INTJ, N, NUM, PART, PRO, V, V.CVB, V.MSDR, V.PTCP
Person	0, 1, 2, 3, 4, EXCL, INCL, OBV, PRX
Polarity	NEG, POS
Politeness	AVOID, COL, FOREG, FORM, FORM.ELEV, FORM.HUMB, HIGH, HIGH.ELEV, HIGH.SUPR, INFM, LIT, LOW, POL
Possession	ALN, NALN, PSSD, PSSPNO
Switch-Reference	CN_R_MN, DS, DSADV, LOG, OR, SEQMA, SIMMA, SS, SSADV
Tense	1DAY, FUT, HOD, IMMED, PRS, PST, RCT, RMT
Valency	DITR, IMPRS, INTR, TR
Voice	ACFOC, ACT, AGFOC, ANTIP, APPL, BFOC, CAUS, CFOC, DIR, IFOC, INV, LFOC, MID, PASS, PFOC, RECP, REFL

Table 1. Dimensions of meaning and their features, both sorted alphabetically

The universal morphological feature schema is composed of a set of features that represent semantic "atoms" that are never decomposed into more fine-grained meanings in any natural language. This ensures that the meanings of all morphemes are able to be represented either through single features or through multiple features in combination.

The purpose of the universal morphological feature schema strongly influences its relationship to linguistic theory. The features instantiated in the schema occupy an intermediate position between being universal categories and comparative concepts, in the terminology coined by Haspelmath [27, pp. 663-667]. Haspelmath defines a universal category as one that is universally available for any language, may be psychologically 'real,' and is used for both description/analysis and comparison while a comparative concept is explicitly defined by typologists, is not claimed to be 'real' to speakers in any sense, and is used only for the purpose of language comparison.

Because the purpose of the schema is to allow broad cross-linguistic morphological analysis that ensures semantic equality between morphemes in one language and morphemes, wordforms, or phrases in another, its features are assumed to be possibly applicable to any language. In this sense, features are universal categories. However, like comparative concepts, the features of the universal schema are not presumed to be 'real' to speakers in any sense.

Like both universal categories and comparative concepts, each feature retains a consistent meaning across languages such that every time a feature is associated with a morpheme, that morpheme necessarily bears the meaning captured by that feature (even though that morpheme may bear other meanings and serve other functions as well). This emphasis on semantic consistency across languages prevents categories from being mistakenly equated, as in the dative case example in Haspelmath [27, p. 665], which highlights the problems with establishing cross-linguistic equivalence on the basis of terminology alone.

2.2 Constructing the Schema

The first step in constructing the universal feature schema was to identify the dimensions of meaning (e.g., case, number, tense, mood, etc.) that are expressed by overt, affixal inflectional morphology in the world's languages. These were identified by surveying the linguistic typology literature on parts of speech and then identifying the kinds of inflectional morphology that are typically associated with each part of speech. In total, 23 dimensions of meaning were identified.

For each dimension, we determined the finest-grained distinctions in meaning that were made within that dimension by a natural language by surveying the literature in linguistic typology. That is, we identified which meanings were "atomic" and were never further decomposed in any language. The reduction of the feature set in the universal schema to only those features whose meanings are as basic as possible minimizes the number of features and allows more complex meanings to be represented by combining features from the same dimension. In addition to these basic features, some higher-level features that represented common cross-linguistic groupings were also included. For example, features such as indicative (IND) and subjunctive (SBJV) represent groupings of multiple basic modality features which nevertheless seem to occur in multiple languages and show very similar usage patterns across those languages [41]. These can be viewed as 'cover features' in which backing off to more basic features remains an option.

Each dimension has an underlying semantic basis that is used to define the features subsumed by that dimension. To determine the underlying semantic basis for each dimension, the linguistic typology and descriptive linguistic theory literature were surveyed for explanations that were descriptively-oriented and offered precise definitions for observed basic distinctions. A simple example is the dimension of number, whose eight features are defined according to a straightforward quantificational scale of the number of entities. The following section presents the schema in detail, describing the semantic basis of each dimension and listing its features.

Because this is the first instantiation of this particular schema, it is likely not yet fully exhaustive and the authors invite input on dimensions or features that should be considered for inclusion. Future work will focus on the possible inclusion of additional features, especially from other known frameworks such as GOLD [24]. Many of the features from the Universal Dependencies Project [51] and the Leipzig Glossing Rules [12] are already integrated into the schema.

2.3 Dimensions of Meaning Encoded by Inflectional Morphology

The semantic bases of the dimensions of meaning that are encoded by inflectional morphology are discussed approximately according to the part of speech with which the dimension is conventionally associated. After the parts of speech themselves, the following dimensions are discussed: (verbs:) Tense, aspect, Aktionsart, mood, voice, evidentiality, switch-reference, person, (nouns:) number, gender, case, animacy, possession, information structure, politeness, (adjectives:) comparison, (pronouns:) deixis. This order is purely expositional: Dimensions of meaning and their features are not formally associated with any particular part of speech.

For reasons of space, we omit discussion of the dimensions of finiteness, interrogativity, and polarity, which exhibit simple binary oppositions, as well as valency and animacy, whose features are typical and defined in the expected way. We also omit discussion of definiteness, which uses features inspired by the the work of Lyons [40, pp. 50, 99, 278]. These dimensions and their features are included in Table 1.

Parts of Speech Croft [16, p. 89] defines the conceptual space in Table 2 for parts of speech. It is the cross-product of the concepts of *object*, *property*, and *action* with the functions of *reference*, *modification*, and *predication*. This conceptual space provides definitions for the following cross-linguistically common parts of speech, which are all captured by features in the universal schema: Nouns (N), adpositions (ADP), adjectives (ADJ), verbs (V), masdars (V.MSDR), participles (V.PTCP), converbs (V.CVB), and adverbs (ADV).

	Reference	*Modification*	*Predication*
Object	object reference: nouns	object modifier: adpositions	object predication: predicate nouns
Property	property reference: substantivized adjectives	property modifier: (attributive) adjectives	property predication: predicate adjectives
Action	action reference: masdars	action modifier: adverbs, participles converbs	action predication: verbs

Table 2. Functionally-motivated conceptual space defining basic parts of speech, adapted from Croft [16, p. 89]

Masdars, participles, and converbs are distinct parts of speech which are nonfinite and derived productively from verbs [26, pp. 4-5]. Masdars (verbal nouns) refer to the action of a verb, such as *running* in *the running of the race*. Participles can be property modifiers when they function like adjectives, and action modifiers when they function like adverbs. Both adverbs and converbs (i.e., verbal adverbs) modify the action expressed by the verb.

In addition to these parts of speech, the following parts of speech are included based on their use in the Universal Dependencies Project [51], which provides an annotation system for approximately 30 languages: Pronoun (PRO), determiner (DET), auxiliary (AUX), conjunction (CONJ), numeral (NUM), particle (PART), and interjection (INTJ). In addition to these, articles (ART), classifiers (CLF), and complementizers (COMP) were given features based on their inclusion in the Leipzig Glossing Rules [12].

Tense Tense and aspect are defined according to the framework in [32], which uses the concepts of Time of Utterance (TU, '|'), Topic Time (TT, '[ ]'), and Situation Time (TSit, '{ }') to define tense and aspect categories. Topic Time (TT) and Situation Time (TSit) are conceived as spans while Time of Utterance (TU) is a single point. By defining tense and aspect categories solely in terms of the ordering of these spans and TU, tense and aspect categories can be defined in a language-independent way that facilitates cross-linguistic comparison.

TU is the time at which a speaker makes an utterance, and topic time is the time about which the claim in the utterance is meant to hold true. TSit is the time in which the state of affairs described by the speaker actually holds true. Tense is the relationship of TU to TT while aspect is the relationship of TT to TSit. The three core tenses are defined schematically in (1-3). To simplify the examples of tense, imperfective aspect is always used (i.e., TT is within TSit).

(1) Past tense (PST): TT precedes TU
————————[————]———|——
'The book was lying on the table.'

(2) Present tense (PRS): TU is within TT
————————[——|——]————————
'The book is lying on the table.'

(3) Future tense (FUT): TU precedes TT
——|————[————]————————
'The book will be lying on the table.'

Some languages further distinguish tense categories by morphologically marking the temporal distance between TU and TT. For example, Bamileke-Ngyemboon (Bantu) distinguishes four levels of temporal distance symmetrically in the past and future, such that for the past there is hodiernal (earlier today; HOD), hesternal (yesterday; 1DAY), recent past (in the last few days; RCT), and remote (RMT) past while for the future there is later today, tomorrow, within the next few days (recent future), and farther ahead yet (remote future) [10, p. 96]. Bamileke-

Dschang (Bantu) also has a symmetrical system, but adds an 'immediate' step (IMMED) indicating 'just now' or 'coming up in a moment' [10, p. 97].

Aspect Aspect indicates the relationship between the time for which a claim is made (TT) and the time for which a situation was objectively true (TSit). The aspects that can be defined by relating TSit and TT are: Imperfective (IPFV), perfective (PFV), perfect (PRF), progressive (PROG), and prospective (PROSP). The iterative (ITER) and habitual (HAB) aspects, sometimes categorized as Aktionsarten, can also be defined this way, but require more than one TSit.

Before defining each category, it is necessary to differentiate 1-state and 2-state verbs. A 1-state verb is a verb like 'sleep,' which lexically encodes only one state (symbolized as '—'). In a 2-state verb, the verb lexically encodes a source state (SS, symbolized as '———') and a target state (TS, symbolized as '++++++'). The verb 'leave' is a 2-state verb, since it is impossible to leave without going through a transition of being somewhere (the source state) and then being gone from that place (the target state).

In the schematic definitions of aspect categories that follow, time of utterance is fixed in the diagrams at a point toward the end of the target state such that all examples are past tense. Note that English does not clearly morphologically distinguish perfective, perfect, and prospective aspects. This complicates translation of the diagrams, but demonstrates their utility in establishing language-independent definitions of these categories.

(4) Imperfective aspect: TT fully within TSit
———————{—[—++]++}+++++++++|++
'She was leaving.'

(5) Progressive aspect: TT is located only within the source state of TSit
———————{—[——]++++++}++++++|++
'She was leaving.'

(6) Perfective aspect: Partial TT overlap with source state or target state
———————[—{—]—++++}+++++++++|++
'She was about to leave.' (source state overlap)
———————{——++[++}++]++++++|++
'She had left.' (target state overlap)

(7) Perfect aspect: TT is located exclusively within the target state of TSit
———————{———++[++++]}++++++|++
'She left. / She has left.'

(8) Prospective aspect: TT is located before TSit
——[——]—{———++++++}++++++|++
'She was going to leave. / She was about to leave.'

(9) Iterative aspect: Multiple instances of the same TSit occur fully within a bounded TT
.....[.....{—+++}$_1$.....{—+++}$_2$.....{—+++}$_n$.....].....|.....
'He used to leave often.'

(10) Habitual aspect: Infinite instances of the same TSit occur fully within an unbounded TT
$[\infty....\{—+++\}_{n}.....\{—+++\}_{n+1}.....|.....\{—+++\}_{n+\infty}....\infty]$
'He (always) leaves early every morning.'

Aktionsart Aktionsart refers to the "inherent temporal features" of a verb [32, pp. 29-31], and is a grammatical means of encoding how the action described by a verb unfolds in reality. We include the distinctions defined by Cable [6], Comrie [8], and Vendler [52]. The features that apply to verbs are Stative (STAT), Eventive/Dynamic (DYN), Telic (TEL), Achievement (ACH), Punctual (PCT), Accomplishment (ACCMP), Durative (DUR), Atelic (ATEL), Semelfactive (SEMEL), and Activity (ACTY).

Mood Grammatical mood is the morphological marking of modality, which "is concerned with the status of the proposition that describes the event" [41, p. 1]. The morphological marking of modality tends to group primary categories of modality into larger superordinate categories. The indicative (IND) and subjunctive (SBJV), realis (REAL) and irrealis (IRR), and Australian non-purposive (AUNPRP) and purposive (AUPRP) moods are superordinate groupings of primary modalities. Each pairs of groupings has a set of core uses that can be reduced to an opposition between indicating information that is asserted as truth and indicating information that is not asserted as truth [41, p. 3]. These superordinate categories are encoded as features for the reasons stated in §2.2.

Basic modality categories that are typically captured by overt morphology include, first, the imperative-jussive modality (IMP). Imperative-jussive statements express a command for an actor to do something. Imperatives typically refer to commands to a second person actor while jussives command a first person plural or third person actor [41, p. 81]. No case was found in which imperative and jussive modalities were contrasted overtly on the same person. Other basic modality categories express varying speculative attitudes, including likely (LKLY), potential (POT), and unlikely or surprising. The Papuan language Dani contrasts the realis, likely, and potential moods overtly [41, p. 162]. Related to the potential mood is the permissive (PERM) mood, which indicates 'may' in the sense of having permission. A number of Balkan languages, including Bulgarian, mark the admirative modality (ADM), which expresses surprise, doubt, or irony [p. 11]. The North American isolate Tonkawa explicitly marks the opposite of speculative, the intentive (INTEN), which expressed "(definitely) will, going to" [p. 82]. Languages such as Tiwi (isolate; Australia) mark the obligative (OBLIG) modality overtly to indicate "must, have to" [p. 75]. Similar to the obligative, the debitive modality (DEB), "ought to, should," is marked overtly in Tamil [p. 27]. The general purposive (PURP) modality indicates 'in order to, for the purpose of.' The conditional mood, familiar from Spanish, expresses "would (if certain conditions held)," and the simulative, which occurs in Caddo, expresses hypothetical action in the sense of "as if X-ing" [41, p. 178]. Finally, the optative or desiderative modality (OPT) marks that an actor wants an action to occur.

Voice Voice is the dimension of meaning that "expresses relations between a predicate [typically a verb] and a set of nominal positions - or their referents - in a clause or other structure" [30]. Klaiman [p. 2] defines three types of grammatical voice: Derived, basic, and pragmatic voice systems.

Derived voice includes two voice categories familiar from Indo-European languages, active (ACT) and passive (PASS). In ergative-absolutive languages, an ergative subject is demoted to an absolutive subject in what is termed an antipassive (ANTIP) construction [30, p. 230]. Derived voice can also include middle voice (MID) in languages like Sanskrit, but middle voice is more often part of basic voice systems (as in Modern Fula), in which voice is captured by lexical items, which have an inherent voice associated with them [30, p. 26].

Pragmatic voice systems include what have been called direct-inverse systems, common in North American languages, as well as complex voicing systems in Austronesian languages. In languages with direct-inverse voice systems (e.g., Plains Cree), arguments are ranked according to a salience hierarchy, such as 1 > 2 > 3 > non-human animate > inanimate. When the most "salient" argument of the verb functions as the subject, the verb may be marked with a direct voice (DIR) morpheme [30, p. 230]. The inverse voice (INV) marks the argument of the verb that is lower in the hierarchy when it functions as the subject. When the arguments of the verb are at equal ranks, they are marked as either proximate or obviative, as described in §2.3 (Person).

In Austronesian voice systems, a different voice is used to focus nouns occupying different semantic roles [30, p. 247]. A voice marker that simultaneously marks the semantic role of the focused noun is used on the verb and the overt marker of the semantic role is replaced by a morpheme that marks both the semantic role and its status as focused. The Austronesian language that makes the most distinctions in semantic role marking in its voice system is Iloko (Ilocano). The semantic roles it marks are given dedicated features in the universal schema since they are used by other Austronesian languages. Those roles are: Agent (AGFOC), patient (PFOC), location (LFOC), beneficiary (BFOC), accompanier (ACFOC), instrument (IFOC), and conveyed (CFOC; either by actual motion or in a linguistic sense, as by a speech act) [45, pp. 336-338].

Finally, valency-changing morphology is categorized with voice because it alters the argument structure of a sentence. Reflexives (REFL) direct action back onto a subject, while reciprocals (RECP) indicate that with a plural subject, non-identical participants perform the action of the verb on each other. Causatives (CAUS) indicate that an action was forced to occur, and may introduce an argument indicating the actant that was forced to perform the action. Applicative morphemes (APPL) increase the number of oblique arguments (that is, arguments other than the subject or object) that are selected by the predicate [42].

Evidentiality Evidentiality is the morphological marking of a speaker's source of information [1]. The universal morphological feature schema follows Aikhenvald [1] in viewing evidentiality as a separate category from mood and modality. Although categories of evidentiality may entail certain modalities (such

as hearsay or reported information evidentials entailing irrealis or subjunctive moods), evidentiality is a distinct category that encodes only the source of the information that a speaker is conveying in a proposition.

The unique evidential categories proposed as features here are based on Aikhenvald's typology [1, pp. 26-60]. Those features are, in approximate order of directness of evidence: Firsthand (FH), direct (DRCT), sensory (SEN), non-visual sensory (NVSEN), auditory (AUD), non-firsthand (NFH), quotative (QUOT), reported (RPRT), hearsay (HRSY), inferred (INFER), and assumed (ASSUM). The degree to which these categories could be reduced using a deeper featural analysis requires further research.

Switch-Reference Switch-reference is an anaphoric linkage between clauses that disambiguates the reference of subjects and other NPs [48, p. 1]. Switch-reference is a fully grammaticalized phenomenon in some languages and can occur when the reference of subjects or other NPs is already fully disambiguated. Switch-reference marking is concentrated in languages of North America (notably in the Southwest, Great Basin, and coastal Northern California), Australia, Papua New Guinea, and the Bantu languages of Africa [48, p. 5].

A basic overt distinction in many switch-reference systems is between same subject (SS) and different subject (DS) [48, pp. 3-4]. In addition to this basic distinction, a third underspecified category, open reference (OR) marking, which signals "indifference as to the referential relation between the two [NPs] rather than specified non-identity" [48, p. 34]. In addition, some West African languages have what have been called "logophoric" systems in which pronouns are explicitly coreferential (or logophoric; LOG) with a pronoun in a previous clause [48, pp. 50-56].

More complex switch-reference systems necessitate additional features, which, due to space limitations, are not described here, but are included in the summary of the schema. Note that CN_R_MN is a feature template used to signal switch-reference marking between NPs in any argument position (as must be used for, e.g., Warlpiri) [48, p. 25]. When expanded, these template features bring the total feature count above 212.

Person The conventional person categories that are encoded on verbs in most languages include first person (1), second person (2), and third person (3). Apart from these common distinctions, some languages also distinguish other categories of person, including zero (0) and fourth person (4), and each conventional person category is sometimes subdivided further. The Santa Ana dialect of Keres distinguishes all four of these categories [20, pp. 75-76].

Zero person, which occurs in Finnish, describes an underspecified third person, as with English 'one,' that refers to any human actor [34, p. 209]. Fourth person is used to describe an otherwise third-person referent that is distinguished via switch-reference (e.g., in Navajo "disjoint reference across clauses" [56, p. 108]) or obviation status [7, pp. 306-307].

The first person plural ('we') is divided into inclusive (INCL), i.e., including the addressee, or exclusive (EXCL), i.e., excluding the addressee. When two or more third person arguments are at the same level of the salience hierarchy in a language with a direct-inverse voice system, one argument is usually overtly marked as proximate (PRX) and the other as obviative (OBV).

Number The dimension of number is relevant for multiple parts of speech and is one of the most frequent agreement features. Each feature is defined with respect to a quantificational scale of the number of entities indicated. The range of number distinctions on nouns is most extensive, with less common categories like "greater paucal" expressed in a small number of languages on nouns, but never on verbs.

The number categories found on nouns include singular (SG), plural (PL), dual (DU), trial (TRI), paucal (PAUC), greater paucal (GPAUC), and so-called inverse number (INVN) [14]. Sursurunga (Austronesian) contrasts all these, except inverse, on nouns [14, pp. 25-30].

In inverse number systems, such as that of Kiowa [14, pp. 159-161], nouns have a default number that indicates the number with which they are "expected" to occur. For example, if 'child' is by default singular and 'tree' is by default plural, then inverse number marking would make 'child' plural and 'tree' singular, inverting the number value of the noun.

Gender Gender is a grammatical category that includes both conventional gender from European languages like Spanish and German, and systems with more than three categories that are typically described as noun class systems.

Because gender can be assigned according to semantic, morphological, phonological, or lexical criteria, creating an underlying conceptual-semantic space for defining gender features is of limited utility. In addition, gender categories rarely map neatly across languages, with differences in gender assignment even where semantic criteria primarily determine gender. This schema therefore treats gender as an open-class feature. The working strategy for limiting feature proliferation is to encode features for gender categories that are shared across languages within a linguistic family or stock in order to capture identical gender category definitions and gender assignments that result from common ancestry. Results presented in Table 3a. offer evidence that this is an effective strategy, given the level of agreement in gender features within a family. The features masculine (MASC), feminine (FEM), and neuter (NEUT) are motivated by many Indo-European languages. To capture the eight possible Nakh-Daghestanian noun classes, the features NAKH1, NAKH2, etc. are used, and to capture the Bantu noun classes, of which 25 are estimated to have existed in Proto-Bantu [21, p. 272], the features BANTU1, BANTU1A, BANTU2, etc. are used.

Case "Case is a system of marking dependent nouns for the type of relationship they bear to their heads" [3, p. 1]. The types of overt case that are encountered

in the world's languages can be divided into three types: 1) core case, 2) local case, and 3) other types of case [3].

Core case is also known as 'non-local,' 'nuclear,' or 'grammatical' case [3, 13], and indicates the role of a syntactic argument as subject, object, or indirect object. The specific core cases vary according to the syntactic alignment that a given language uses and can be defined in terms of three standard "meta-arguments," S (subject of an intransitive verb), A (subject of a transitive verb), and P (object of a transitive verb). Nominative-accusative languages use the nominative case (NOM) to mark S and A and the accuative (ACC) to indicate P. Ergative-absolutive languages use the ergative case (ERG) to indicate A and absolutive (ABS) to indicate S and P. In 'tripartite' languages that fully differentiate S, A, and P, the S-only nominative (NOMS) indicates only S.

Non-core, non-local cases (type 3) express non-core argument relations and non-spatial relations. The dative case (DAT) marks the indirect object, and its functions are sometimes divided into two distinct cases, the benefactive (BEN) for marking the beneficiary of an action and the purposive (PRP) for marking the reason or purpose for an action [3, pp. 144-145]. The genitive (GEN) and relative (REL) cases both mark a possessor, with relative also marking the core A role [p. 151]. The partitive case (PRT) marks a noun as partially affected by an action [p. 153]. The instrumental case (INS) marks the means by which an action is done, and sometimes marks accompaniment, which can be marked distinctly with the comitative case (COM) [p. 156]. The vocative case (VOC) marks direct address [pp. 4-5]. In comparative constructions, the standard of comparison (e.g. 'taller *than X*') can be explicitly marked with the comparative case (COMPV) when the comparison is unequal and with the equative case (EQTV; e.g., 'as much *as X*') when the comparison is equal. The formal case (FRML) marks "in the capacity of, as," and the aversive case (AVR), common in Australian languages, indicates something that is to be feared or avoided. Also common in Australian languages are the privative/"abessive" case (PRIV) indicating without or a lack or something and its counterpart, the proprietive case (PROPR), which indicates the quality of having something [3, p. 156].

The local cases express spatial relationships that are typically expressed by adpositions in English (and in the majority of the world's languages) [44, p. 24]. The types of local case morphemes include place, distal, motion, and 'aspect' morphemes, as shown by Radkevich [**?**].[1] The place morphemes indicate orientation to a very precise degree [p. 29]. The Nakh-Daghestanian languages Tabassaran and Tsez contain the largest number of place morphemes, which include separate morphemes, encoded in the schema as features, for "among (INTER), at (AT), behind (POST), in (IN), near (CIRC), near/in front of (ANTE), next to (APUD), on (ON), on (horizontal; ONHR), on (vertical; ONVR)," and "under (SUB)" [44, 13]. Only one morpheme (and feature) indicates distal (REM). The motion category is composed of only three possible parameters, namely

[1] The local case morphemes can be organized within each category through the use of abstract features that are more general than the feature labels employed in the schema.

essive (static location; ESS), allative (motion toward; ALL), and ablative (motion away; ABL) [44, pp. 34-36]. The 'aspect' category is an elaboration of the motion category, and includes four parameters, namely approximative (APPRX), terminative (TERM), prolative/translative (PROL), and versative (VERS) [pp. 37, 53-55]. The approximative indicates motion toward, but not reaching, a goal, while the terminative indicates that motion "as far as," or "up to" the goal. The versative indicates motion in the direction of a goal, without indication of whether it is reached, and the prolative/translative indicates motion "along, across," or "through" something.

Animacy To the extent that animacy is a grammatically separate category from person, individuation, and agency, it encompasses only four principal categories: Human (HUM), non-human (NHUM), animate (ANIM), and inanimate (INAN) [11, p. 185]. Animacy is not encoded by dedicated overt morphemes in any language, but can still be isolated as an independent parameter that has overt morphological effects. Animacy conditions the realization of accusative case in Russian, with animate masculine nouns taking a form identical to the genitive and inanimate masculine nouns taking a form identical to the nominative [58, p. 48].

Possession Some languages, including Turkish and certain Quechua languages, use overt affixal morphology to mark characteristics of the possessor directly on a possessed noun or to encode the type of possession. The simplest type of marking on the possessed noun marks no characteristics of the possessor, but simply encodes the quality of being possessed (PSSD). This feature occurs in Hausa, Wolof, and in the construct state in Semitic languages [15].

The grammatical characteristics of the possessor that are marked in languages of the world include person, clusivity, number, gender, and politeness. For example, Huallaga Quechua marks person, clusivity, and number [53, pp. 54-55]. Turkish marks person, number, and formality [23, p. 66], and Arabic marks person, number (including dual), and gender (masculine and feminine) [46, p. 301]. The features used to capture these morphemes contain the prefix PSS-, followed by a number indicating person (1-3), S, D, or P for number, I or E for clusivity, M or F for gender, and INFM or FORM for politeness. For example, possession by a second person singular masculine possessor is marked with the feature PSS2SM. This feature is schematized as PSSPNO ('possession-person-number-other').

Finally, many languages (such as Kpelle [Mande]), distinguish alienable possession (ALN), in which ownership can change, from inalienable possession (NALN), in which ownhership is considered to be inherent. For example, Kpelle marks possession by a first person singular possessor distinctly in 'my house' (*ŋa pɛrɛi*) from 'my arm' (*m-pôlu*) [54, p. 279].

Information Structure Information structure is a component of grammar that formally expresses "the pragmatic structuring of a proposition in a discourse" [35, p. 5]. More concretely, information structure directly encodes which parts

of a proposition are asserted by the speaker (the focus; FOC) and which are presupposed or otherwise not asserted (the topic; TOP; ibid., pp. 5-6).

The topic signals what the sentence is about. Lambrecht [35, p. 131] defines the topic more specifically as "expressing information which is relevant to [a referent in the proposition] and which increases the addressee's knowledge of this referent." The focus signals information that is not presupposed by the addressee [35, p. 213]. The information marked by the focus forms the core of the proposition's assertion, and typically includes the part of the proposition that is unpredictable or new to the listener (ibid.).

Politeness Politeness is the dimension of meaning that expresses social status relationships between the speaker, addressee, third parties, or the setting in which a speech act occurs [9, 5]. Politeness/honorific systems can indicate relationships along four axes: 1) The speaker-referent axis, 2) the speaker-addressee axis, 3) the speaker-bystander axis, and 4) the speaker-setting axis [9, 5].

Levinson [36, p. 90] writes that with honorifics along the speaker-referent axis, "respect can only be conveyed by referring to the 'target' of the respect" and that "the familiar *tu*/*vous* type of distinction in singular pronouns of address ... is really a referent honorific system, where the referent happens to be the addressee." The T-V distinction encodes the informal (INFM) and formal (FORM) distinction. Data from Japanese motivate positing two sublevels of the formal level. Japanese uses one set of referent honorifics in a speech style called *sonkeigo* to elevate the referent (FORM.ELEV) and a distinct set of referent honorific forms in a speech style called *kenjōgo* to lower the speaker's status (FORM.HUMB), thereby raising the referent's status by comparison [55, pp. 41-43].

In speaker-addressee honorific systems, politeness is conveyed by word choice itself, not just by terms that refer to the addressee. Japanese and Javanese use these systems, and in each, the distinction is between a polite form (POL) that conveys respect and a plain form that does not.

Features are defined for speaker-bystander honorific systems, as occur in Dyirbal (Pama-Nyungan) and Pohnpeian (Austronesian) [36, pp. 90-91], for example, and for the speaker-setting axis (or register), but are not described here due to space limitations.

Comparison Comparison and gradation can be expressed through overt affixal morphology [18]. The comparative (CMPR), such as English *-er*, relates two objects such that one exceeds the other in exhibiting some quality (ibid.). The superlative (SPRL) relates any number of objects such that one exceeds all the others. This is specifically the relative (RL) superlative, such as that expressed by English *-est*. Another type of superlative, the absolute (AB) superlative, expresses a meaning like "very" or "to a great extent," and is used in Latin, for example [18]. Equative constructions are comparative constructions in which the compared entities exhibit a quality to an equal extent. The adjective itself can be marked as conveying equality (EQT), as in Estonian and Indonesian [18].

Deixis Deictic features, primarily spatial, are used to differentiate third-person pronouns and demonstrative pronouns, especially in languages where these categories overlap [2, pp. 134-135]. Contrasts can be established according to distance, verticality, reference point, and visibility. The maximal distance distinction occurs in Basque, which contrasts proximate (PROX), medial (MED), and remote (REMT) entities [28, pp. 123, 150]. The maximal number of verticality distinctions occurred in the original Lak (Nakh-Daghestanian) pronoun system, which contrasted remote pronouns that encoded being below (BEL), at the same level as (EVEN), or above (ABV) the speaker [22, p. 304]. The maximal reference point distinction occurs in Hausa, which contrasts a pronoun with proximity to the first person (speaker; REF1), to the second person (addressee; REF2), and to neither ('distal'; NOREF) [2, p. 145]. Finally, the maximal visibility distinction occurs in Yupik (Eskimo-Aleut), which distinguishes visible (VIS) from invisible (NVIS), and further subdivides visible elements into those that are 'extended,' i.e., spread out and moving (e.g., the ocean), and those that are 'restricted,' i.e., in sight and stationary [4]. More research into distinctions in the visibility domain is required before positing features beyond VIS and NVIS.

3 Enabling Projection-Based Approaches to Fine-Grained Morphological Tagging

A primary motivation for richly annotating inflectional morphology in a consistent, universally-applicable way is that it enables direct comparison (and even translation) across languages. In this section, we examine variability in the use of inflectional morphological features across languages. Understanding this variability is central to evaluating the viability of simple projection-based approaches (such as those developed by [57, 29, 50, 19]) to fine-grained part-of-speech tagging (i.e., morphological tagging), particularly of underspecified languages.

Some languages, such as English, lack significant surface morphology, so many semantic distinctions must be discovered through contextual analysis. For example, English lacks overt indicators of politeness on verbs, whereas many other languages (e.g., Japanese, Spanish) express it directly through inflectional morphology. If we align the underspecified English word to its foreign counterparts (using standard tools from machine translation), they could provide a consensus label for unspecified semantic values. These consensus-derived labels could be used to generate training data for monolingual semantic tagging algorithms, without the need for costly human annotation effort. The quality of the labels would depend on the tendency of foreign languages to consistently realize inflectional features.

The following sections present a method of measuring cross-linguistic variability in inflectional morphology in order to assess the validity of projection-based approaches to tagging.

3.1 Bible Alignments

We examined cross-linguistic variability in inflectional morphology by comparing which morphological features were expressed across multiple translations of the same meaning. First, we use a set of locations in the New Testament portion of the New International Version (NIV) of the English Bible as 'pivots.' A location is described by a (verse, position) pair and constitutes a context-specific word-meaning combination. All (and only) nominal and verbal words in the NIV New Testament were used as pivots.

For each pivot, we found all single-word foreign translations using verse-level alignments obtained from the Berkeley aligner [39] on the 1169 bibles from `http://paralleltext.info/data/all/`. It was possible for a given pivot to be translated into the same foreign language multiple times, if multiple versions of the Bible were available in that language.

Foreign words were then linked to universal morphological feature representations in our schema via lookup in a database of richly annotated data from Wiktionary.[2] The database contained inflected wordforms from 1,078,020 unique lemmas across the 179 languages represented in Wiktionary's English edition. For further details on the extraction of Wiktionary data and mapping those data to features in the universal morphological feature schema, see Sylak-Glassman, Kirov, Yarowsky, and Que [49].

To avoid ambiguity, only words with a single unique feature vector were used. A total of 1,683,086 translations were able to be mapped this way. Overall, these covered 47 unique languages across 18 language families (e.g., Romance, Celtic, Slavic, Germanic, Uralic, Quechuan, etc.). Family affiliation was determined by manually correcting output from Ethnologue [37]. These mappings made it possible to quantify the level of agreement in feature value for each dimension of meaning across different translations of the same pivot. See Figure 1 for an example in which pairwise agreement may be measured between a Spanish and Russian translation of the same English pivot word. This example also shows how an underspecified English wordform can be labeled with additional morphological features via consensus of its non-English counterparts.

3.2 Results and Discussion

As an indicator of cross-linguistic consistency, Table 3a. describes the average percentage of translation pairs (e.g., see Figure 1) that agree on a particular feature across available pivots.[3] For a particular dimension, only pairs of translations that both specify a non-null feature value were ever compared. The table shows the average pairwise agreement for each dimension across all translations, the average when comparisons are limited to translations from *different* language

[2] `http://www.wiktionary.org`

[3] Some disagreement in the data will be due to errors in our Wiktionary data, or the automated Bible alignment. We do not discuss these sources of noise in this paper, but they should affect all measurements in a uniform way, and thus do not preclude the comparisons we make.

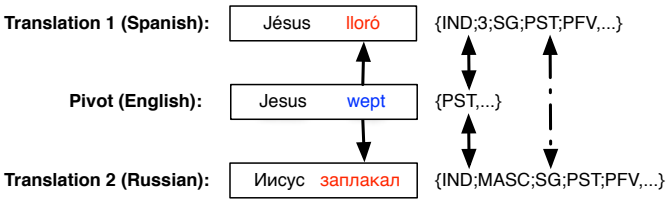

Fig. 1. Pairwise agreement of multiple translations (Spanish and Russian) of the same (English) pivot location. Note that the pivot word in this case, *wept*, only has the PST (past tense) feature overtly specified in English. However, we can assign it other labels including SG and PFV through a consensus of the available translations.

families, the average when comparisons are limited to the same language family, and the average when comparisons are limited to the same language (i.e., only between different Bible versions).

The results indicate that within-language variability is very low. This is an upper bound measuring variability due to translators' linguistic choices, rather than true differences in cross-language feature realization. There is more variability within language families, but the overall drop in agreement is small. This suggests that consensus-based labeling of a target language would be very effective if parallel data from genealogically-related languages were available. Surprisingly, this is true for gender, which, aside from animate nouns with natural masculine or feminine gender, is often assumed to be assigned arbitrarily or according to non-semantic principles [17]. Our data indicate that gender assignment tends to be preserved as related languages diverge from a common proto-language.

Even if we only have parallel text from a set of mutually unrelated languages, the different families column in Table 3a. suggests that we may still rely on a solid consensus for many features. Gender, and presumably other arbitrarily-assigned features do show significant drop in agreement across unrelated languages.

Nominal case shows especially poor agreement cross-linguistically. There are a number of possible reasons for this. First, no core case features will agree between languages with different syntactic alignment systems. Second, languages sometimes assign morphological case in idiosyncratic ways. For example, Russian uses instrumental case not only to denote an implement, but also to mark the time of day and season of the year that an action takes place [43]. These linguistic sources of disagreement, combined with a larger overall set of possible labels for the case feature, predict a lower base rate of agreement.

While pairwise agreement statistics provide a general idea of the feasibility of cross-linguistic projection depending on the similarity of available translation languages to the target, they are not a direct evaluation of the accuracy of consensus-based labels. Since we do not currently have hand-labeled gold-standard data with which to perform such an evaluation, we offer three approximations, shown in Table 3b. The held-out column shows the probability that,

(a.)

Dimension	Overall	Different Family	Same Family	Same Language
Case	0.45	0.23	0.77	0.91
Gender	0.75	0.39	0.87	0.96
Mood	0.89	0.82	0.95	0.99
Number	0.79	0.74	0.88	0.96
Part of Speech	0.74	0.73	0.85	0.94
Person	0.87	0.82	0.93	0.97
Politeness	0.98	0.84	0.99	1.00
Tense	0.73	0.66	0.82	0.95
Voice	0.95	0.83	0.99	0.99
Average	0.79	0.67	0.89	0.96

(b.)

Dimension	Held-Out	Albanian	Latin
Case	0.50	0.57	0.81
Gender	0.76	0.74	0.44
Mood	0.91	N/A	0.96
Number	0.83	0.83	0.85
Part of Speech	0.83	0.86	0.59
Tense	0.79	0.84	0.65
Voice	0.95	N/A	0.84
Average	0.80	0.77	0.73

Table 3. Table (a.) summarizes cross-linguistic agreement for each feature dimension. The 'overall' results correspond to pairwise agreement across all available translations. The 'different family' column shows pairwise agreement among only translations from different language families. The 'same family' and 'same language' columns show pairwise agreement only between translations from the same family, and the same language, respectively. Table (b.) summarizes cross-linguistic projection accuracy for each feature dimension. The 'held-out' column indicates the probability that a held-out translation for an English pivot will match the consensus of the remaining translations. The Albanian and Latin columns indicate the accuracy of consensus compared to gold-standard Albanian and Latin feature labels provided by automatic feature-extraction from Wiktionary.

across all translations of a given pivot, the feature values of a single held-out translation match the consensus values from the remaining translations (i.e., each held-out translation acts as proxy for a gold-standard). The rows in the Albanian and Latin columns show the result of using Albanian and Latin Bibles as a source of pivot locations, and treating our automatically-derived Wiktionary data for these languages as a gold-standard.[4] Albanian is an especially interesting case. Because it is an isolate within the larger Indo-European family, no highly genealogically similar languages were available in our dataset. This simulates the labeling of an unknown new language.

[4] When comparing Albanian and Latin pivots to the consensus of their translations, no Albanian and Latin translations were used. Using only cross-language consensus prevents unfair advantage from self-similarity.

Overall, the results indicate that an approach based on consensus would be effective for assigning feature labels to wordforms. This is especially true if data from languages within the same family are available. For many feature dimensions, even cross-family labels would be useful, especially in low-resource environments where a large gold-standard training set is otherwise unavailable. The high levels of cross-linguistic agreement, particularly for non-arbitrary semantic distinctions, would not be possible if our feature schema could not be consistently applied to multiple, potentially unrelated languages.

4 Conclusion

The universal morphological feature schema presented here incorporates findings from linguistic typology to provide a cross-linguistically applicable method of describing inflectional features in a universalized framework. It greatly expands the coverage of inflectional morphological features beyond previous frameworks and at the same time offers a substantive hypothesis on the dimensions of meaning and which distinctions within them are encoded by inflectional morphology in the world's languages.

The schema offers many potential benefits for NLP and machine translation by facilitating direct meaning-to-meaning translations across language pairs, regardless of form-related differences. We demonstrated that Wiktionary forms, when annotated according to our schema, were very likely to agree along the dimensions of meaning expressed by inflectional morphology when they were aligned to the same pivot words by automatic machine translation tools. This cross-linguistic consistency supports the viability of consensus-based multilingual projection of fine-grained morphological features to an underspecified target language (e.g., tagging formality levels in English even though they are not expressed by the native inflectional system) when parallel text is available.

References

1. Aikhenvald, A.Y.: Evidentiality. Oxford University Press, Oxford (2004)
2. Bhat, D.N.S.: Pronouns. Oxford University Press, Oxford (2004)
3. Blake, B.J.: Case. Cambridge University Press, Cambridge (2001)
4. Bliss, H., Ritter, E.: Developing a Database of Personal and Demonstrative Pronoun Paradigms: Conceptual and Technical Challenges. In: Proceedings of the IRCS Workshop on Linguistic Databases. IRCS, Philadelphia (2001)
5. Brown, P., Levinson, S.C.: Politeness: Some Universals in Language Usage. Cambridge University Press, Cambridge (1987).
6. Cable, S.: Tense, Aspect and Aktionsart, `http://people.umass.edu/scable/PNWSeminar/handouts/Tense/Tense-Background.pdf`
7. Chelliah, S.L., de Reuse, W.J.: Handbook of Descriptive Linguistic Fieldwork. Springer, Dordrecht (2011)
8. Comrie, B.: Aspect: An Introduction to the Study of Verbal Aspect and Related Problems. Cambridge University Press, Cambridge (1976)

9. Comrie, B.: Linguistic Politeness Axes: Speaker-Addressee, Speaker-Referent, Speaker-Bystander. In: Pragmatics Microfiche 1.7 (1976)
10. Comrie, B.: Tense. Cambridge University Press, Cambridge (1985)
11. Comrie, B.: Language Universals and Linguistic Typology. Basil Blackwell (1989)
12. Comrie, B., Haspelmath, M., Bickel, B.: Leipzig Glossing Rules, https://www.eva.mpg.de/lingua/resources/glossing-rules.php
13. Comrie, B., Polinsky, M.: The Great Daghestanian Case Hoax. In: Case, Typology, and Grammar: In Honor of Barry J. Blake, pp. 95–114. John Benjamins (1998)
14. Corbett, G.: Number. Cambridge University Press: Cambridge (2000).
15. Creissels, D.: Construct Forms of Nouns in African Languages. In: Proceedings of the Conference on Language Documentation and Linguistic Theory 2, pp. 73–82. SOAS, London (2009)
16. Croft, W.: Parts of Speech as Language Universals and as Language-Particular Categories. In: Approaches to the Typology of Word Classes, pp. 65–102. Mouton de Gruyter, New York (2000)
17. Cucerzan, S., Yarowsky, D.: Minimally Supervised Induction of Grammatical Gender. In: Proceedings of HLT-NAACL, pp. 40–47. ACL, Stroudsburg, PA (2003)
18. Cuzzolin, P., Lehmann, C.: Comparison and Gradation. In: Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung / An International Handbook on Inflection and Word-Formation, pp. 1212–1220. Mouton de Gruyter, Berlin (2004)
19. Das, D., Petrov, S.: Unsupervised Part-of-Speech Tagging with Bilingual Graph-Based Projections. In: Proceedings of the Association for Computational Linguistics, pp. 600–609. ACL, Stroudsburg, PA (2011)
20. Davis, I.: The Language of Santa Ana Pueblo. In: Anthropological Papers, Numbers 68-74, Bureau of American Ethnology, Bulletin 191, pp. 53–190. Smithsonian Institution, Washington, DC (1964)
21. Demuth, K.: Bantu Noun Classes: Loanword and Acquisition Evidence of Semantic Productivity. In: Classification Systems, pp. 270–292. Cambridge University Press, Cambridge (2000)
22. Friedman, V.: Lak. In: Encyclopedia of Language and Linguistics, pp. 303–305. Elsevier (2006)
23. Göksel, A., Kerslake, C.: Turkish: A Comprehensive Grammar. Routledge (2005)
24. General Ontology for Linguistic Description (GOLD). http://linguistics-ontology.org/
25. Greenough, J.: Allen and Greenough's New Latin Grammar. Dover Publications, Newburyport, MA (2013)
26. Haspelmath, M.: The Converb as a Cross-Linguistically Valid Category. In: Converbs in Cross-Linguistic Perspective, pp. 1–56. Mouton de Gruyter, Berlin (1995)
27. Haspelmath, M.: Comparative Concepts and Descriptive Categories in Crosslinguistic Studies. Language 8(3), 663–687 (2010)
28. Hualde, J.I., Ortiz de Urbina, J.: A Grammar of Basque. Mouton de Gruyter (2003)
29. Hwa, R., Resnik, P., Weinberg, A., Cabezas, C., Kolak, O.: Bootstrapping Parsers via Syntactic Projection Across Parallel Texts. Natural Language Engineering 11, 311–325 (2005)
30. Klaiman, M.H.: Grammatical Voice. Cambridge University Press (1991)
31. Klein, H.G.: Tempus, Aspekt, Aktionsart. Max Niemeyer Verlag, Tübingen (1974)
32. Klein, W.: Time in Language. Routledge, New York (1994)
33. Klein, W.: A Time-Relational Analysis of Russian Aspect. Language 71, 669–695 (1995)

34. Laitinen, L.: Zero Person in Finnish: A Grammatical Resource for Construing Human Reference. In: Grammar from the Human Perspective: Case, Space and Person in Finnish, pp. 209–232. John Benjamins, Amsterdam (2006)
35. Lambrecht, K.: Information Structure and Sentence Form. Cambridge University Press, Cambridge (1994)
36. Levinson, S.C.: Pragmatics. Cambridge University Press, Cambridge (1983)
37. Lewis, M.P., Simons, G.F., Fennig, C.D.: Ethnologue: Languages of the World (18th Edition), http://www.ethnologue.com. SIL International, Dallas, TX.
38. Lewis, W., Xia, F.: Developing ODIN: A Multilingual Repository of Annotated Language Data for Hundreds of the Worlds Languages. Literary and Linguistic Computing 25, 303–319 (2010)
39. Liang, P., Taskar, B., Klein, D.: Alignment by Agreement. In: Proceedings of HLT-NAACL, pp. 104–111. ACL, Stroudsburg, PA (2006)
40. Lyons, C.: Definiteness. Cambridge University Press, Cambridge (1999)
41. Palmer, F.R.: Mood and Modality. Cambridge University Press: Cambridge (2001)
42. Polinsky, M.: Applicative Constructions, http://wals.info/chapter/109
43. Pulkina, I., Zaxava-Nekrasova, E.: Russian: A Practical Grammar with Exercises. Russky Yazyk Publishers, Moscow (1992)
44. Radkevich, N.: On Location: The Structure of Case and Adpositions. University of Connecticut, Storrs (2010)
45. Rubino, C.: Iloko. In: The Austronesian Languages of Asia and Madagascar, pp. 326–349. Routledge, London (2005)
46. Ryding, K.C.: A Reference Grammar of Modern Standard Arabic. Cambridge University Press, Cambridge (2005)
47. Sagot, B., Walther, G. Implementing a Formal Model of Inflectional Morphology. In: Mahlow, C., Piotrowski, M. (eds.) Systems and Frameworks for Computational Morphology, pp. 115–134. Springer, Heidelberg (2013)
48. Stirling, L.: Switch-Reference and Discourse Representation. Cambridge University Press, Cambridge (1993)
49. Sylak-Glassman, J., Kirov, C., Yarowsky, D., Que, R.: A Language-Independent Feature Schema for Inflectional Morphology. In: Proceedings of the ACL. ACL, Stroudsburg, PA (to appear)
50. Täckström, O., Das, D., Petrov, S., McDonald, R., Nivre, J.: Token and Type Constraints for Cross-Lingual Part-of-Speech Tagging. Transactions of the Association for Computational Linguistics 1, 1–12 (2013)
51. Universal Dependencies, http://universaldependencies.github.io/docs/
52. Vendler, Z.: Verbs and Times. The Philosophical Review 66, pp. 143–160 (1957)
53. Weber, D.J.: A Grammar of Huallaga (Huanuco) Quechua. University of California Press, Berkeley (1989)
54. Welmers, W.E.: African Language Structures. University of California Press, Berkeley (1973)
55. Wenger, J.R.: Some Universals of Honorific Language with Special Reference to Japanese. University of Arizona, Tucson (1982)
56. Willie, M.: Navajo Pronouns and Obviation. University of Arizona, Tucson (1991)
57. Yarowsky, D., Ngai, G., Wicentowski, R.: Inducing Multilingual Text Analysis Tools via Robust Projection across Aligned Corpora. In: Proceedings of the First International Conference on Human Language Technology Research, pp. 1–8. ACL, Stroudsburg, PA (2001)
58. Yamamoto, M.: Animacy and Reference. John Benjamins, Amsterdam (1999)