



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2020

---

## **Category clustering: A probabilistic bias in the morphology of verbal agreement marking**

Mansfield, John ; Stoll, Sabine ; Bickel, Balthasar

**Abstract:** Recent research has revealed several languages (e.g. Chintang, Rarámuri, Tagalog, Murrinh-patha) that challenge the general expectation of strict sequential ordering in morphological structure. However, it has remained unclear whether these languages exhibit random placement of affixes or whether there are some underlying probabilistic principles that predict their placement. Here we address this question for verbal agreement markers and hypothesize a probabilistic universal of category clustering, with two effects: (i) markers in paradigmatic opposition tend to be placed in the same morphological position ('paradigmatic alignment'; Crysmann Bonami 2016); (ii) morphological positions tend to be categorically uniform ('featural coherence'; Stump 2001). We first show in a corpus study that category clustering drives the distribution of agreement prefixes in speakers' production of Chintang, a language where prefix placement is not constrained by any categorical rules of sequential ordering. We then show in a typological study that the same principle also shapes the evolution of morphological structure: although exceptions are attested, paradigms are much more likely to obey rather than to violate the principle. Category clustering is therefore a good candidate for a universal force shaping the structure and use of language, potentially due to benefits in processing and learning.\*

DOI: <https://doi.org/10.1353/lan.2020.0021>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-188214>

Journal Article

Published Version

Originally published at:

Mansfield, John; Stoll, Sabine; Bickel, Balthasar (2020). Category clustering: A probabilistic bias in the morphology of verbal agreement marking. *Language. Journal of the Linguistic Society of America*, 96(2):255-293.

DOI: <https://doi.org/10.1353/lan.2020.0021>



PROJECT MUSE®

---

Category clustering: A probabilistic bias in the morphology  
of verbal agreement marking

John Mansfield, Sabine Stoll, Balthasar Bickel

Language, Volume 96, Number 2, June 2020, pp. 255-293 (Article)

Published by Linguistic Society of America

DOI: <https://doi.org/10.1353/lan.2020.0021>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/757629>

# CATEGORY CLUSTERING: A PROBABILISTIC BIAS IN THE MORPHOLOGY OF VERBAL AGREEMENT MARKING

JOHN MANSFIELD

SABINE STOLL

BALTHASAR BICKEL

*University of Melbourne*

*University of Zurich*

*University of Zurich*

Recent research has revealed several languages (e.g. Chintang, Rarámuri, Tagalog, Murrinhpatha) that challenge the general expectation of strict sequential ordering in morphological structure. However, it has remained unclear whether these languages exhibit random placement of affixes or whether there are some underlying probabilistic principles that predict their placement. Here we address this question for verbal agreement markers and hypothesize a probabilistic universal of CATEGORY CLUSTERING, with two effects: (i) markers in paradigmatic opposition tend to be placed in the same morphological position ('paradigmatic alignment'; Crysmann & Bonami 2016); (ii) morphological positions tend to be categorically uniform ('featural coherence'; Stump 2001). We first show in a corpus study that category clustering drives the distribution of agreement prefixes in speakers' production of Chintang, a language where prefix placement is not constrained by any categorical rules of sequential ordering. We then show in a typological study that the same principle also shapes the evolution of morphological structure: although exceptions are attested, paradigms are much more likely to obey rather than to violate the principle. Category clustering is therefore a good candidate for a universal force shaping the structure and use of language, potentially due to benefits in processing and learning.\*

*Keywords:* morphology, linguistic typology, verbal agreement, morphotactics, linguistic complexity, language evolution

**1. INTRODUCTION.** A hallmark of morphology has been the assumption that morphemes are rigidly ordered (e.g. Anderson 1992:261), but several languages defy this expectation by allowing free placement of elements inside a word.<sup>1</sup> Chintang (Sino-Tibetan; Bickel et al. 2007), Mari (Uralic; Luutonen 1997), Rarámuri (Uto-Aztecan; Caballero 2010), Tagalog (Austronesian; Ryan 2010), Murrinhpatha (Southern Daly; Mansfield 2015), and a few others (Rice 2011) allow affix placement that is variable and not regulated by any formal or semantic factors. The data in 1 illustrate the phenomenon in Chintang, where all logically possible arrangements of prefixes are equally grammatical, with no known effect on meaning (Bickel et al. 2007).

\* This article benefited from insightful comments by Rebecca Defina, Roger Levy, David Nash, Rachel Nordlinger, Sebastian Sauppe, Robert Schikowski, and three anonymous referees. We also received helpful comments after presentations at the Surrey Morphology Group, the Australian Linguistics Society conference in 2017, the Societas Linguistica Europaea in 2019, and the Association for Linguistic Typology in 2019. JM's work on this article was supported by the ARC Centre of Excellence for the Dynamics of Language (Project ID: CE140100041) and an Endeavour Fellowship from the Australian Government Department of Education and Training. SST's work was supported by the project 'Acquisition processes in maximally diverse languages: Min(d)ing the ambient languages (ACQDIV)', which has received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7-2007-2013) (Grant agreement No. 615988; PI Sabine Stoll). BB's work was supported by Swiss National Science Foundation Sinergia Grant No. CRSIII\_160739.

<sup>1</sup> By 'word' we mean here combinations of morphemes whose selection constraints ban inserting phrasal material (Bickel & Zúñiga 2017), following other approaches that use interruptibility criteria (e.g. Nercesian 2014, Green & Morrison 2016).

- (1) a. u-kha-ma-cop-yokt-e  
 3NSG.A-1NSG.P-NEG-see-NEG-PST<sup>2</sup>  
 b. u-ma-kha-cop-yokt-e  
 3NSG.A-NEG-1NSG.P-see-NEG-PST  
 c. kha-u-ma-cop-yokt-e  
 1NSG.P-3NSG.A-NEG-see-NEG-PST  
 ... *etc.*

all: 'They didn't see us.'

(Bickel et al. 2007:44)

Free variation in placement is even more widely attested in clitics, that is, dependent elements that are less tightly integrated into grammatical words (e.g. Bickel et al. 2007 on Swiss German, Diesing et al. 2009 on Serbian, Good & Yu 2005 on Turkish, Harris 2002 on Udi, Schwenter & Torres Cacoullos 2014 on Spanish).

While progress has been made in formal models of free placement (e.g. Crysmann & Bonami 2016, Ryan 2010), an unresolved question is whether the production of variable placement is subject to systematic probabilistic principles, or whether it is solely governed by chance and perhaps processing-internal factors such as priming or lexical access. In other words, given a choice of morphotactic variants, do speakers select forms that conform to some principles at a rate higher than chance? Or are variants selected with equal probability, or perhaps probabilities that are specific to each element or context of occurrence?

Here, we propose that affix placement is subject to a probabilistic universal of CATEGORY CLUSTERING, with two effects.

- (2) CATEGORY CLUSTERING: Morphological categories tend to cluster in positions, that is,  
 a. markers of the same category tend to be expressed in the same morphological position, and  
 b. morphological positions tend to be filled by markers of the same category.

The two effects are probabilistic versions of principles that are independently motivated in morphological theory: 2a corresponds to PARADIGMATIC ALIGNMENT (Crysmann & Bonami 2016), and 2b to FEATURAL COHERENCE (Stump 2001:20).

We propose the trend for category clustering as a cognitive bias that shapes linguistic structure worldwide, similar in spirit to other such biases that have been found in various aspects of grammar and phonology (e.g. Bickel et al. 2015, Bresnan et al. 2001, Christiansen & Chater 2008, Culbertson et al. 2012, Dediu et al. 2017, Hawkins 1994, 2014, Himmelmann 2014, Kemmerer 2012, MacDonald 2013, Napoli et al. 2014, Seifart et al. 2018, Widmer et al. 2017). As such, we expect its effects to be detectable both in patterns of language production by individual speakers and in worldwide distributions as the result of language change over multiple generations of speakers (Bickel 2015).

We first probe the evidence in language production. For this, we focus on the case of Chintang where speakers have a choice between different ways of ordering markers in sequence. We predict that even when the order of markers is not determined by grammatical rule, it is more likely to comply with category clustering than not. We test this prediction in a naturalistic setting by analyzing corpus data, controlling for competing

<sup>2</sup> In morphological glosses for this article we use full caps for agreement roles and small caps for all other grammatical categories, as follows: A: most agent-like argument of multiargument verbs, ABIL: ability, CAUS: causative, DU: dual, EXCL: exclusive, F: feminine, FUT: future, G: goal or recipient argument of three-argument verbs, HOD: hodiernal, IMP: imperative, INCL: inclusive, IND: indicative, INV: inverse, M: masculine, NEG: negative, NFUT: nonfuture, NMLZ: nominalizer, NPST: nonpast, NSG: nonsingular, P: most patient-like argument of multiargument verbs, PAUC: paucal, PL: plural, PRF: perfect, PST: past, RED: reduplicant, REL: relative, S: sole argument, SG: singular, TEL: telic, TRANS: transitive.

factors such as idiosyncrasy of speakers or lexical items and persistence effects of simply reusing the same order that last occurred in the discourse.

In a second study we examine the effects of the bias on global distributions. For this we sample morphological data from languages that have fixed ordering in word forms. We predict that these languages have evolved in such a way that their attested orders are more likely to show category clustering than not. In other words, when morphological categories evolve through grammaticalization and reanalysis, we expect such processes to keep entire categories together in specific positions and to gradually erode any earlier marking of the same categories that might be present in different positions. We test this prediction against a typological data from the AUTOTYP database of grammatical markers (Bickel, Nichols, et al. 2017).

In both studies we focus on agreement markers on the verb (i.e. various kinds of cross-references to verbal arguments) because this is where we have the richest set of data.

Evidence for category clustering in morphology has implications for theories of the morphology vs. syntax division. If morphological markers tend to be placed according to the categories (features) they express—for example, agreement markers placed according to syntactic role (subject, object, etc.)—this suggests that morphological sequencing follows principles similar to those of syntax, where positions are equally linked to categories. This favors theories that assume a single, unified system for morphology and syntax alike (e.g. Embick & Noyer 2007, Halle & Marantz 1993). By contrast, if languages deviate from category clustering, with idiosyncratic rules that place, say, first-person subject expressions into different positions from third-person subject expressions, this favors theories that posit an autonomously organized morphology component (e.g. Anderson 1992, Inkelas 1993, Stump 1997, 2001). Any apparent similarities between morphology and syntax would then be merely historical residues of the phrase structures from which morphology evolves (Givón 1971).

While this debate has been mostly framed as a categorical choice between theoretical options, we assess category clustering in a probabilistic way. In other words, we explore the impact of categorical clustering vs. idiosyncrasy on affix ordering as a typological variable (Bickel & Nichols 2007, Simpson & Withgott 1986), treating the distinctiveness of morphology from syntax as an evolving property of languages rather than a theoretical choice in the formal architecture of grammar.

Below we first introduce the principle of category clustering in more detail and situate it with respect to the literature on affix placement (§2). We then test our prediction on variable prefix sequences in Chintang (§3) and on the global distribution of fixed-position affixes (§4). In §5 we discuss the implications of our findings for morphological theory, and hypothesize an explanation of the clustering bias in terms of its benefits for processing and learning.

**2. CATEGORY CLUSTERING IN THEORETICAL PERSPECTIVE.** Morphological theories generally expect category clustering. This expectation is explicitly captured by the two principles of paradigmatic alignment (Crysmann & Bonami 2016) and featural coherence (Stump 2001).

**2.1. PARADIGMATIC ALIGNMENT.** Taking feature values as a starting point, Crysmann and Bonami (2016:317) observe that canonical affix systems cluster by the principle of paradigmatic alignment (also see Anderson 1992:131). We paraphrase their technical definition as follows.

- (3) **PARADIGMATIC ALIGNMENT:** Affixes that encode contrasting values of the same grammatical feature appear in the same position relative to the stem and to other affixes.

Paradigmatic alignment states only that affixes of the same category occur in the same position, but it does not at the same time require that a given position contain only affixes of the same category. This additional requirement is captured instead by featural coherence (see §2.2 below).

An example of paradigmatic alignment can be seen in Amele verb affixation (Trans-New Guinea; Roberts 1987), where the morphosyntactic features G (goal) agreement, A (agent-like) agreement, and tense-aspect-mood (TAM) each have consistent suffix positions.

- (4) a. *ija ho faj-h-ig-en*  
 1SG work show-2SG.G-1SG.A-FUT  
 ‘I will show it to you.’  
 b. *uqa ahul geli-te-i-a*  
 3SG coconut scrape-1SG.G-3SG.A-HOD.PST  
 ‘She scraped the coconut for me.’  
 c. *hu-t-ag-a*  
 come-1SG.G-2SG.A-IMP  
 ‘Come to me!’

(Amele; Roberts 1987:281)

While paradigmatic alignment is canonical, it is not difficult to find instances of ‘misalignment’, that is, where affixes of the same category appear in different positions. One example is in Fula (Niger-Congo; Arnott 1970), where A and P (patient-like) agreement suffixes appear in different relative orders depending on person (and number) values, as in 5. Placement may also vary in relation to the stem, as in Khaling (Tibeto-Burman; Jacques et al. 2012), where most S (intransitive subject) agreement affixes are suffixes, but the second-person singular S exponent is a prefix, as in 6.

- (5) a. *mball-u-moo-mi*  
 help-REL.PST-3SG.P-1SG.A  
 ‘I helped him.’  
 b. *mball-u-daa-mo*  
 help-REL.PST-2SG.A-3SG.P  
 ‘You helped him.’

(Fula; Stump 2001:151)

- (6) a. *mu-ŋa*  
 be-1SG.S  
 b. *?i-mu*  
 2SG.S-be  
 c. *mu-nu*  
 be-3SG.S

(Khaling; Jacques et al. 2012:1124)

Another type of alignment violation is caused by multiple exponence, where grammatical categories are spread over multiple positions (Harris 2017). For example, in Murrinhpatha verbs (Southern Daly; Mansfield 2019, Nordlinger 2015) the number and gender of the A role is expressed jointly in three morphological positions. The prefix position encodes A person and number, the first suffix position further encodes A number, and a later suffix position encodes A number and gender.

- (7) *pu-mam-ka-ŋime*  
 3PL.A-say.NFUT-PAUC.A-PAUC.F.A  
 ‘They (paucal, fem.) said.’

(Murrinhpatha; Mansfield 2019:112, 142)

Since agreement markers often fuse features specifying person and number with features specifying roles, paradigmatic alignment in one feature often comes at the expense of misalignment in the other feature. In the examples of agreement markers above, we have focused on (mis)alignment of markers by role. But in some circum-

stances, agreement markers may instead paradigmatically align by grammatical person and other referential categories (Nichols 1992), or by complex constellations of arguments (Witzlack-Makarevich et al. 2016). For example, in Anindilyakwa (Macro-Gunwinyguan; van Egmond 2012), some agreement prefixes in realis mood are paradigmatically aligned into two positions, one for first and second person, followed by one for third person, independently of the roles expressed.

- (8) a. (nə)nge-nə-rrəŋka  
1SG.A-3M.P-see.PST  
'I saw him.'
- b. ngə-nə-rrəŋka  
1SG.P-3M.A-see.PST  
'He saw me.'
- c. kərrə-nga-rrəŋka  
2PL.A-3F.P-see.PST  
'You (pl.) saw her.'
- d. kərr-angə-rrəŋka  
2PL.P-3F.A-see.PST  
'She saw you (pl).'

(Anindilyakwa; van Egmond 2012:140–41)

**2.2. FEATURAL COHERENCE.** The other aspect of category clustering, featural coherence, takes morphological positions, rather than affixes, as its starting point. Paraphrasing Stump's (2001:20) technical formulation (see also Good 2016:55), we define this as follows.

- (9) FEATURAL COHERENCE: For any given morphological position, affixes placed in that position should express the same feature(s), and by implication, affixes expressing different features should be placed in different positions.

Featural coherence can apply to multiple features simultaneously if a position systematically bundles features such as person and number. We refer to affixes being of the same 'category' as a shorthand for such bundling, so that featural coherence is a matter of positions hosting a single affix category.

Featural coherence often follows naturally from paradigmatic alignment, since if affixes of the same category are consistently allocated to a particular position, then the position might be expected to host a uniform category. This is the case for the Amele example cited above, where not only do G markers, A markers, and TAM categories have consistent placement, but they also each have a different consistent placement and therefore do not occupy the same positions. However, there are two ways in which paradigmatic alignment may be satisfied while featural coherence is not. First, portmanteau affixes may encode multiple agreement roles in a single affix position, such that each role satisfies paradigmatic alignment, but the position now hosts two different roles at the same time, violating featural coherence. This is again exemplified in Anindilyakwa. While we saw above that A and P may be independently expressed for some person values, transitive verbs involving combinations of first and second person encode both roles in a single portmanteau prefix.

- (10) a. yirra-rrəŋka  
1>2-see.PST  
'I saw you.'
- b. yə-rrəŋka  
2>1-see.PST  
'You saw me.'

(Anindilyakwa; van Egmond 2012:140–41)

The second way that paradigmatic alignment can be satisfied while featural coherence is violated is when affixes of different categories appear to ‘compete’ for the same morphological position. An example is the Algonquian prefix position where A and P markers compete: second-person agreement is selected over first-person agreement, regardless of role (which is differentiated by suffixes). We illustrate with an example from Wôpanâak, the Massachusetts language for which the phenomenon was first described (von Humboldt 1836:189ff.).

- (11) a. ku:w-adchan-eh  
 2-keep-2SG>1SG (prefix marking 2.A)  
 ‘You keep me.’  
 b. nu-ttunn-uk  
 1-say-INV (prefix marking 1.P)  
 ‘He said to me.’ (Goddard & Bragdon 1988:520ff., Fermino 2000)

In competition phenomena of this type, agreement markers paradigmatically align in the prefix position, but the position is not featurally coherent with respect to role.

The Anindilyakwa data in 7 above also show featural incoherence (each position contains markers of A or P), but in addition, they violate paradigmatic alignment (not all A markers cluster in the same position).

To summarize paradigmatic alignment and featural coherence, Table 1 schematically illustrates some affixal phenomena that satisfy (‘+’) or violate (‘-’) each principle.  $A_1$ ,  $A_2$  and  $P_1$ ,  $P_2$  are sets of same-category affixes from which at most one affix can be selected (with the subscript indexing person),  $A_1 > P_2$  is a portmanteau affix, and  $A_{1\alpha}$ ,  $A_{1\beta}$  is a category expressed by a combination of two affixes (i.e. multiple exponence). Disjunctive sets like  $(A_1 | A_2)$  share a morphological position, and positions are separated by hyphens. Our schema here illustrates each phenomenon separately, while actual languages may combine elements of more than one phenomenon (e.g. portmanteau and multiple exponence).

	+ FEATURAL COHERENCE	- FEATURAL COHERENCE
+ PARADIGMATIC ALIGNMENT	STEM-( $A_1   A_2$ )-( $P_1   P_2$ )	STEM-( $A_1   A_2   P_1   P_2$ ) STEM-( $A_1 > P_2   A_2 > P_1$ )
- PARADIGMATIC ALIGNMENT	STEM- $A_1$ -( $P_1   P_2$ )- $A_2$ STEM-( $A_{1\alpha}   A_2$ )-( $P_1   P_2$ )- $A_{1\beta}$	STEM-( $A_1   P_1$ )-( $A_2   P_2$ ) STEM-( $A_1 > P_{2\alpha}$ )-( $P_{2\beta}$ )

TABLE 1. Schematic representation of clustering satisfaction and violation in A vs. P agreement categories.

**2.3. THEORETICAL IMPLICATIONS OF CATEGORY CLUSTERING.** Category clustering and its violation are highly consequential for morphological theory, and in particular for debates about the morphology/syntax interface. Morphological theory is riven by disagreement about whether word structure should be seen as a subdomain of syntax (e.g. Embick & Noyer 2007, Halle & Marantz 1993) or whether there is an entirely autonomous system of morphology (e.g. Anderson 1992, Stump 2001). All theories agree that there is at least a partially systematic ‘grammar of words’, but at the same time, compared to phrase structure, word structure exhibits more arbitrary and idiosyncratic combinatorics. Rules of affix placement make this tension between systematicity and idiosyncrasy particularly prominent (Manova & Aronoff 2010) and have therefore played a key role in morphological theory (Bickel & Nichols 2007, Hyman 2003, Rice 2000, Stump 1997).

Nonclustered, misaligned, and idiosyncratic affix placement suggests that general syntactic principles fail to account naturally for morphological positioning (Inkelas 1993, Simpson & Withgott 1986, Stump 1997). This provides room for a relatively autonomous morphology in the overall architecture of grammar. In some instances,



nonclustered affixation may be explained as an epiphenomenon of other features, for example, phonological patterns (Hyman 2003:270, Kim 2010, Rice 2011). However, it remains doubtful that all instances can be reanalyzed in this way (Bickel & Nichols 2007, Paster 2009, Witzlack-Makarevich et al. 2016).

Conversely, category clustering may be interpreted as evidence for a close relationship between morphology and syntax. Category clustering is a syntax-like pattern, since in most theories of phrase structure, node positions are identified with broad categories, rather than particular lexical items (Good 2016:55).<sup>3</sup> Thus, in phrase structures such as [Det Adj N] or [NP [V NP]], positions collect and cluster consistent categories of lexical items. Indeed, lexical categories are to a large extent DEFINED by the positions they can occur in, and syntax is commonly taken to be driven by these categories. Category clustering is therefore tacitly assumed as a principle of syntax. Consistent with this assumption, 86% of languages in Dryer's (2013) database have a dominant constituent order based on agent-like vs. patient-like roles. Most of the remaining languages show orders that are chiefly driven by information-structure categories (focus, topic, etc.) or syntactic categories such as subordination or auxiliiation. Indeed, while it remains an important target of research, some striking instances of free word order have been shown to be sensitive to information structure (e.g. Warlpiri; Simpson 2007, Simpson & Mushin 2008). Free prefix ordering in Chintang is more radical because it is not affected by information structure or any pragmatic constraint (Bickel et al. 2007). However, the question remains of whether this free ordering shows probabilistic tendencies toward category clustering, potentially reducing the gap with syntax.

Independently of whether theories posit a fundamental morphology vs. syntax division, virtually all assume that category clustering is a default, while violations of clustering require more complex placement specifications. Thus, even in theories like PARADIGM FUNCTION MORPHOLOGY (PFM; Stump 2001), which treats morphology as largely autonomous from syntax, category clustering is taken to be the default insofar as realization rules are organized into 'rule blocks'. Surface clustering in morphological positions is modeled in PFM as affixes belonging to a common rule block. These are sets of rules from which exactly one is applied, that being the rule that matches the largest subset of features that need to be expressed and enter word formation (instantiating what is known as Paninian competition). The resolution of competing rules in each block depends upon blocks being composed of contrasting values of the same syntactic feature(s)—that is, they are featurally coherent (Stump 2001:20). For example, if a block were to contain one rule selecting TENSE: PST and another selecting AGR: 1SG, there would be no way of resolving the competition for an input bundle matching both of these feature values. Now, since rule blocks must be featurally coherent and the competing realization rules apply at the same point in the derivation, category clustering is the natural outcome of Paninian competition.<sup>4</sup> Rule blocks can also account for affixes on opposite sides of the stem (as in the Khaling example 4), which may be prefix and suffix exponences in the same block.<sup>5</sup> But other forms of misalignment require either a

<sup>3</sup> This is orthogonal to the more specific question of whether the hierarchical structure of syntax is mirrored by the linear positions found in affix positioning. There is substantial research on this latter question, for example, with regard to tense and aspect ordering (Baker 1985, Bybee 1985, Foley & Van Valin 1984, Julien 2002), derivational affixes (Hyman 2003, Rice 2000, Stump 1997), and agreement and case markers (Bickel & Nichols 2007).

<sup>4</sup> This follows only when rule blocks contain more than one rule. If each block contains only a single rule, this (trivially) results in featural coherence, but it does not result in clustering. However, single-rule blocks are only the limiting case of blocks and not their theoretical rationale.

<sup>5</sup> Crysmann and Bonami (2016:339) argue that this situation is rare in language, but should be relatively common under PFM. They propose this as an argument for preferring left-to-right morphological placement, as opposed to stem-centric placement (see also Spencer 2003:639).

stipulative mechanism to reverse the default order of rule blocks for certain feature values, or the splitting up of feature values into multiple rule blocks (cf. Crysmann 2017, Stump 2001:154–56). The capacity for rule blocks to be arbitrarily reordered or multiplied in this way is one dimension in which PFM distinguishes word structure from syntax. In summary, PFM predicts that category clustering should be the norm, while crucially also allowing for violations in its basic architecture.

A more recent theory of autonomous morphology, INFORMATION-BASED MORPHOLOGY (IbM; Crysmann 2017, Crysmann & Bonami 2016), gives more room for nonclustering phenomena. In this approach, inflectional affixes are explicitly specified for position class, rather than being grouped into rule blocks (also see Spencer 2003:640). This means that the competing values of agreement markers (for example) can more freely be realized in distinct positions. Clustering is encoded by shared inheritance of a position class among all exponents of AGR, while nonclustering is encoded by specification on individual exponents (Crysmann & Bonami 2016:359). The IbM approach is therefore well adapted to nonclustering, and although the basic model does not by itself predict a clustering bias, nonclustering requires more complex positional specifications. This explicit encoding of complexity (Sagot & Walther 2011) could provide the basis for deriving a probabilistic clustering bias.

In more syntactically oriented theories of morphology there is naturally a much stronger expectation of category clustering, and violations are less easily accommodated. DISTRIBUTED MORPHOLOGY (DM; Halle & Marantz 1993), for example, proposes that complex words consist of multiple nodes in syntactic structure, with movement, merge, and conditional spell-out operations accounting for the discrepancies between surface word structure and underlying phrase structure (Embick 2015, Embick & Noyer 2001). Since category clustering is expected of syntactic structure, word forms are also expected to exhibit clustering, except for where this is disrupted by postsyntactic operations. Much of the DM literature focuses on morphological movement operations that affect entire syntactic nodes irrespective of their specific feature values (e.g. Embick & Noyer 2001). Such operations maintain category clustering, while allowing for discrepancies between syntactic structure and morphological linearization. Violations of category clustering require alternative mechanisms. Simple prefix/suffix alternation as in Khaling can be dealt with by treating direction of attachment as part of the phonological specification of affixes (Embick & Noyer 2007:317), but in some instances further stipulations are required. For example, in Noyer's (1997:214) analysis of Tamazight Berber, affix placement is generally determined by phrase structure, together with morphological well-formedness constraints. But some affixes have 'free licensing' and are positioned independently of syntax or any other general constraints.

- (12) a. t-dawa  
           3SG.F-cure  
           'She cures.'
- b. dawa-nt  
           cure-3PL.F  
           'They (fem.) cure.'
- c. t-dawa-d  
           2SG.M-cure-2SG.M  
           'You (sg., masc.) cure.'
- d. t-dawa-m  
           2PL.M-want-2PL.M  
           'You (pl., masc.) cure.'

(Noyer 1997:216)

Where misalignment goes beyond simple prefix/suffix alternation, the syntactic approach to morphology may be able to deal with this only by stipulating that some affixes are independent of the general word-structure system.<sup>6</sup> Autonomous-morphology theories such as PFM and IBM are more up-front about nonclustered morphology, since such phenomena are part of their basic architecture.

However, by taking category clustering as a default, theories like PFM risk duplicating principles of syntax in morphology, and such theories must indeed explain why words and phrases share this fundamental property (Stump 2001:21). One potential explanation is that morphology tends to mirror phrase structure merely as a historical residue, even if the synchronic morphology is fundamentally independent from the syntax (e.g. Bybee 1985:38, Crismann & Bonami 2016:334, Ryan 2010:784, Spencer 2003:629, Stump 2001:27). This follows from Givón's (1971) proposal that morphology is a residue of earlier phrase structure. But the extent to which category clustering is indeed preserved during the diachronic transition from syntax to morphology (Anderson 1980) remains an unresolved issue. And if it is preserved, it still needs to be explained why it is preserved despite the many processes that can disrupt clustering in an autonomous morphology, such as erosion, metathesis, or reanalysis.

Languages with grammatically free affix placement add a new perspective to this debate. If there are no placement rules whatsoever, positions cannot be linked to categories (or features, for that matter). However, grammatically free affix placement might still be subject to probabilistic biases of category clustering. If this is the case, the debate about the morphology vs. syntax division is no longer one about theoretical choices, but rather becomes an empirical question about the degree to which languages exhibit category clustering in word and phrase structure. This enlarges the space of expected variation: morphology and syntax may have similar degrees of category clustering and therefore look similar, or they may have different degrees and look different. From this perspective, category clustering is not an intrinsic property of either syntax or morphology, but a more general principle that may shape either domain to different extents.

As a first step in exploring this possibility, we turn to probabilistic approaches to affix placement.

**2.4. PROBABILISTIC APPROACHES.** While there are various probabilistic studies of morphology examining phenomena such as productivity (Hay 2002, Plag & Baayen 2009) and allomorph selection (Ackerman & Malouf 2013), there is very little on probabilities of morphological placement. A few studies have examined variable clitic placement, using logistic regression to model how grammatical and discourse factors influence selection between placement outcomes (Diesing et al. 2009, Schwenter & Torres Cacoullos 2014). With regard to affixes, the most detailed study we know of is on variable placement of aspectual reduplication in Tagalog (Ryan 2010). Tagalog has a 'contemplated' aspect marked by leftward CV: reduplication, and the base for this reduplication may be either the verb stem or one of several prefixes (Schachter & Otanes 1983). Some prefixes are not available as the reduplication base, and some bases have degraded acceptability.

<sup>6</sup> Trommer (2003) argues that agreement affixes are positioned by syntax in concert with optimality-theoretic constraints. Like Noyer, he analyzes an idiosyncratically positioned affix (in this instance in Ancash Quechua) as having a morpheme-specific constraint that sets it apart from the general system. Arregi & Nevins 2012 is a study of agreement placement in Basque, where nonclustering is analyzed in terms of movement rules that target specific feature values (Arregi & Nevins 2012:272).

- (13) ma-ka-pag-pa-sayá  
 ABIL-TEL-TRANS-CAUS-be.happy  
 ‘able to make happy’
- (14) a. ma-ka:ka-pag-pa-sayá  
 ABIL-RED-TEL-TRANS-CAUS-√
- b. ma-ka-pag-pa:pa-sayá  
 ABIL-TEL-TRANS-RED-CAUS-√
- c. ?ma-ka-pag-pa-sa:pa-sayá  
 ABIL-TEL-TRANS-CAUS-RED-√
- d. \*ma-ka-pa:pag-pa-sayá  
 ABIL-TEL-RED-TRANS-CAUS-√  
 ‘will be able to make happy’

(Tagalog; Ryan 2010:764)

Ryan provides a probabilistic analysis of Tagalog reduplication (RED) positioning, based on token frequencies harvested from web search results. He models the variation using HARMONIC GRAMMAR (Smolensky & Legendre 2006), an OPTIMALITY-THEORETIC (OT) model where constraints are ranked on a continuous scale and the probability of an output is a function of summed constraint violation. Each pair of potentially adjacent morphemes is given a ‘coherence’ value (not to be confused with ‘featural coherence’), representing the morphemes’ propensity to be adjacent. The probability of RED appearing in one position or another is determined by the weighting of the morpheme coherence constraints, such that higher probability is generated by positions that accrue lower total constraint violations. Table 2 illustrates an example of a constraint-weighting tableau, in which the preference for 5a over 5b is modeled as a high degree of coherence in the bigram RED-*ka-* compared to a low degree of coherence in the bigram RED-*pa-*. High coherence among other bigrams such as *ka-pag-* disfavors RED placement that would interrupt this bigram, that is, \**ka-RED-pag-* as in 5d.

GEN'D	SCORE	CANDIDATE	RED-ka	pag-RED	ka-pag	pag-pa	pa-RED	RED-pa	ma-RED	ma-ka	pa-√	RED-√
			11.1	10.9	9.6	7.9	7.6	7.2	6.9	6.7	6.5	2.7
a.	74.51%	35.1	ma-RED-ka-pag-pa-√	*			*	*		*		*
b.	25.45%	36.2	ma-ka-pag-RED-pa-√	*		*	*	*				*
c.	0.04%	42.6	ma-ka-pag-pa-RED-√	*	*			*	*		*	
d.	0.00%	56.0	ma-ka-RED-pag-pa-√	*	*	*		*	*	*		*

TABLE 2. Weighted bigram coherence constraints, adapted from Ryan 2010:769.

Ryan’s model is built on constraints that are specific to Tagalog morphology, which he treats as being idiosyncratic and autonomous from syntax (Ryan 2010:778). In the following we propose a different probabilistic approach, modeling affix placement in terms of generic, crosslinguistically relevant principles of category clustering. Furthermore, rather than use a probabilistic OT approach as in Ryan’s study, we use regression models of a type more widely used in various branches of linguistics (e.g. Baayen 2008, Gorman & Johnson 2013, Levshina 2015, Speelman 2014). We test the presence of category clustering first in Chintang (§3) and then in a global database (§4).

Both of our studies focus on agreement markers because it is here where we have the richest data. Also, we focus exclusively on surface positions (sometimes referred to as ‘slots’), as in IBM, rather than rule blocks (PFM) or syntactic derivations (DM). The reason for this choice is that PFM- and DM-style analyses inevitably favor effects of cate-

gory clustering because this is, as we noted, the default assumption in these theories. For example, the distribution of S agreement markers in Khaling in 4 complies with category clustering under a rule-block analysis, but violates category clustering in terms of surface positions. Therefore, data on surface positions are more likely to work against our hypothesis, thus providing a more conservative test for a category clustering bias.

**3. PROBABILISTIC CLUSTERING IN CHINTANG PREFIXES.** Chintang is an Eastern Kiranti (Tibeto-Burman) language spoken by five to six thousand people in the Himalayan foothills of Nepal. Chintang verbs have a complex morphological structure (Bickel et al. 2007, Bickel & Zúñiga 2017, Schikowski 2014, Stoll et al. 2017). In what follows we briefly describe the prefix system, which is the domain of free placement.

**3.1. CHINTANG AFFIX PLACEMENT.** Chintang verb structure is summarized by the regular expression given in 15. ‘Prefix’ and ‘suffix’ refer to elements that can occur with verb stems and are inflectionally required by verb stems; ‘clitic’ refers to various elements attaching at a phrasal level.<sup>7</sup>

(15) (Prefix\*-STEM-Suffix<sup>+</sup>)<sup>+</sup> Clitic\*

While suffix elements have fixed placement, all prefix elements are freely placed relative to one another.

- (16) a. u-kha-ma-cop-yokt-e [= 1]  
           3NSG.A-1NSG.P-NEG-see-NEG-PST  
       b. u-ma-kha-cop-yokt-e  
           3NSG.A-NEG-1NSG.P-see-NEG-PST  
       c. kha-u-ma-cop-yokt-e  
           1NSG.P-3NSG.A-NEG-see-NEG-PST  
       ... *etc.*

all: ‘They didn’t see us.’ (Bickel et al. 2007:44)

The reordering of prefix elements has no effect on meaning, and speakers accept all logically possible orders as equally grammatical. The only constraint is that prefixes need to attach to the left of a (specific kind of) phonological word. This precludes them from occurring inside suffix strings, but allows them to appear, for example, between stems in a compound (Bickel et al. 2007).

Chintang prefixes express negation as well as paradigmatic alternations for agreement with A (most agent-like), P (most patient-like), and S (sole argument of intransitives) roles, with some variation depending on a verb’s lexical valency (Schikowski et al. 2015). Table 3 shows the complete set of prefixes, that is, the set of affixes that are freely ordered. The markers *a-* and *u-* are paradigmatic alternates with different person/number values in S and A agreement, here labeled SUBJ. *a-* encodes SUBJ as second person, underspecified for number; *u-* encodes third nonsingular, or underspecified third person when P is first singular. Other SUBJ agreement values are encoded by suffixes. The markers *ma-* ~ *mai-* both mark first-person nonsingular P agreement, here labeled OBJ. They also encode an inclusive/exclusive distinction, though we collapse this distinction in our coding due to small token numbers (see Supplementary Material 1<sup>8</sup>). The marker *kha-* is another first-person nonsingular OBJ marker, but it socially indexes geographical provenance. Its use is associated by speakers with Sambugañ village,

<sup>7</sup> Chintang also has lexical preverb elements that have a distribution similar to that of prefixes, including variable ordering. Preverbs are not discussed in this study, but for details see Bickel et al. 2007.

<sup>8</sup> The supplementary materials referenced here and throughout can be accessed online at <http://muse.jhu.edu/resolve/96>.

while *ma-* ~ *mai-* are associated with Mulgaũ village, though people from all areas may use either variant. As with SUBJ markers, other OBJ person/number values are expressed by suffixes.

CATEGORY	PREFIX	MEANING	(VILLAGE)
NEG	mai- ~ ma- <sup>a</sup>	NEG	
SUBJ	a-	2.S/A	
	u-	3NSG.S/A; 3.A (if P = 1SG)	
OBJ	kha-	1NSG.P	Sambugaũ
	ma-	1NSG.EXCL.P	} Mulgaũ
	mai-	1NSG.INCL.P	
A>P	na-	3>2	

TABLE 3. Chintang prefixes by grammatical category.

<sup>a</sup> NEG and OBJ markers have homophonous forms, involving free variation of the NEG form, and inclusive/exclusive distinction in the OBJ marker. Our corpus data includes all variants, but in our statistical coding and in the discussion below we normalize these to *mai-* NEG and *ma-* OBJ.

Combinations of these prefixes can occur in any order. Inspection of spontaneous speech data in our corpus confirms that the two SUBJ prefixes can each cooccur with *mai-* NEG in either order.

- (17) a. u-mai-ta-yokt-e  
3NSG.S-NEG-come-NEG-IND.PST  
'They did not come.' (CLLDCh1R04S05.0054)
- b. mai-u-ta-yokt-e  
NEG-3NSG.S-come-NEG-IND.PST  
'They did not come.' (dihi\_khahare.312)
- (18) a. a-mai-apt-th-a-ŋ-ni-ŋ-a  
2.A-NEG-shoot-NEG-IMP-1SG.P-2PL-1SG.P-IMP  
'Don't shoot me!' (Chambak\_int.1213)
- b. mai-a-hid-u-ŋs-u-ce-e-kha  
NEG-2.A-finish-3.P-PRF-3.P-3NSG.P-IND.PST-NMLZ  
'Haven't you finished them?' (kamce\_talk.0020)

The SUBJ prefixes can also cooccur with OBJ prefixes in either order. Of the eight possible bigrams (2 SUBJ × 2 OBJ × 2 orders), seven are found in our data, while the one nonoccurring form *u-ma-* is likely to be an accidental lacuna, as {*u-*, *ma-*} bigrams of either order happen to be rare ( $N = 3$ ). We illustrate {SUBJ, OBJ} bigram variability with the somewhat more frequent *kha-* variant.

- (19) a. a-kha-lus-no  
2.A-1NSG.P-tell-IND.PST  
'You'll tell us.' (ctn\_cut.417)
- b. kha-a-lud-ce-ke  
1NSG.P-2.A-tell.D-IND-PST  
'You'll tell us.' (CLLDCh3R12S03.478)
- (20) a. u-kha-patt-a-ŋs-a-kha  
3NSG.A-1NSG.P-call-PST-PRF-PST-NMLZ  
'They've called us.' (CLLDCh3R02S03.0404)
- b. kha-u-patt-no-go  
1NSG.P-3NSG.A-call-IND.NPST-NMLZ  
'Do they call us?' (CLLDCh2R04S04.32)

The question for our study is, are certain orders preferred beyond chance? And if so, do the placement patterns of prefixes align by grammatical category?

**3.2. CORPUS DATA USED FOR THIS STUDY.** The Chintang corpus (Bickel, Stoll, et al. 2017) comprises 1.3 million words of naturalistic speech by adults and children. The largest part of the corpus consists of spontaneous conversational recordings, complemented by traditional stories, myths, and video-elicited narratives. We focus here on language produced by speakers with adult-like performance in morphology, which is reached after at most five years of age (Stoll et al. 2017).<sup>9</sup>

We extracted from the corpus all sequences of paired prefixes hosted together on a verb ( $N = 621$ ). This includes bigram tokens for all possible combinations of the prefixes in Table 3, given that a verb can host maximally one SUBJ prefix and one OBJ prefix. The vast majority of these tokens ( $N = 603$ ) come from verbs with exactly two prefixes, while the remainder come from verbs with three prefixes, which we count as two pairs A-B, B-C. There are also a number of tokens that involve duplication of the same prefix, or where one or both prefixes are hosted by a secondary verb rather than the main verb stem in compounds. After filtering out these tokens, we have 576 observations of prefix<sub>1</sub>-prefix<sub>2</sub> bigrams, in which we analyze patterns of sequencing. The raw data and scripts used in this study are available on GitHub.<sup>10</sup>

An interesting question that is outside of our present purview is the degree of interspeaker variation. Previous research has shown that individual Chintang speakers use variable prefix orders, which suggests that the variation is present in individuals' mental grammars and is not an artefact of aggregating data from various idiolects (Bickel et al. 2007). Further insight into speaker-level variation will require structured sampling, as the number of tokens per individual is quite small in the corpus (ranging from one to thirty-five).

**3.3. CATEGORY CLUSTERING BY GRAMMATICAL CATEGORY.** We test for probabilistic clustering by investigating whether prefixes of the same grammatical category tend to occur in the same position relative to other prefixes. There are three basic possibilities for how the prefix bigram sequences might be distributed:

- (i) Bigram sequences A-B and B-A have uniform probability; that is, there is no bias toward one sequence or the other;
- (ii) Each bigram type has a bias, such that A-B has a different probability from B-A, but the probabilities of different bigram types are not systematically related to grammatical categories;
- (iii) Each bigram has a bias, and these are systematically related, such that A-B and A-C have a similar probability due to a grammatical similarity between B and C. It is this scenario that exhibits a category clustering effect. If A-B and A-C have similar biases, where B and C are of the same category, this exhibits probabilistic paradigmatic alignment of B and C. Furthermore, if A-B and A-D have different biases, where B and D are of different categories, this exhibits a probabilistic difference among different positions, that is, featural coherence.

Our data suggest that Chintang prefix bigrams follow scenario (iii), that is, they exhibit probabilistic category clustering. First, the bigram data show clear biases toward partic-

<sup>9</sup> Children under five produce very few tokens of the multiply prefixed verbs under discussion, so exclusion of this group removes only 2% of the tokens.

<sup>10</sup> <https://github.com/jbmansfield/Category-Clustering>

ular sequences: the probabilities of A-B and B-A are not uniform. For example, there are 214 tokens combining the SUBJ marker *u-* with the NEG marker *mai-*, and these are distributed with 75 *mai-u-* and 139 *u-mai-* tokens, which would be extremely improbable under a uniform-distribution model (binomial test: 95% CI = [0.58, 0.71], two-sided  $p < 0.001$ ). Crucially, we find that bigrams A-B and A-C have very different probabilities where B and C are of different grammatical categories, but very similar probabilities where B and C mark the same category. For example, we compare the {*u-*, *mai-*} bigrams with {*u-*, *kha-*}, where *kha-* 1NSG.P marks OBJ agreement and *mai-* marks NEG polarity. As shown in Figure 1, each of these bigram types shows a biased distribution, but they are very different biases with respect to the *u-* prefix. *kha-* is to the left of *u-* in 89% of bigram tokens, while *mai-* is to the left in 35% of bigram tokens.

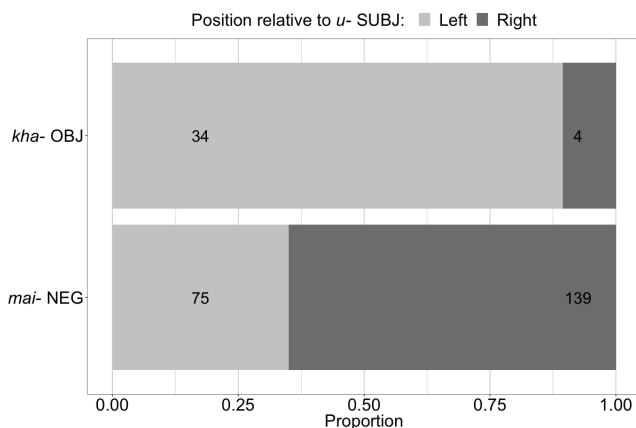


FIGURE 1. Bigrams where the reference prefix *u-* combines with prefixes of different grammatical categories.

In Figure 2 we instead compare bigram types where the same prefix combines with two different prefixes marking the same grammatical category. We take *mai-* NEG combining with either *u-* 3NSG.S/A/3.A or *a-* 2.S/A, that is, two paradigmatically related SUBJ agreement markers. As shown in Fig. 2, the biases of these bigrams are strikingly similar.

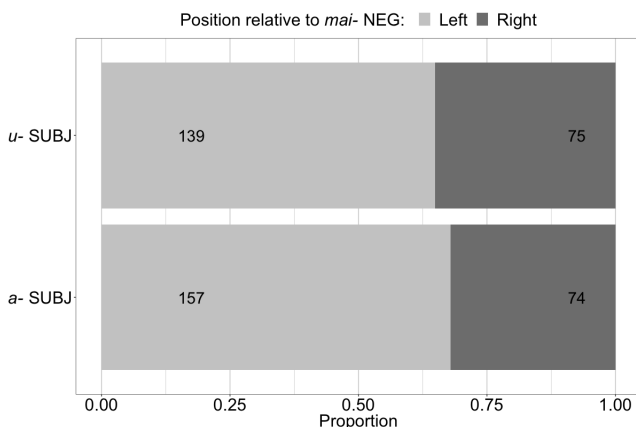


FIGURE 2. Bigrams where the reference prefix *mai-* combines with prefixes of the same grammatical category, SUBJ.



If morphs of the same grammatical category have the same placement biases, then we should be able to treat *u-* and *a-* as a combined reference for comparing bigrams of a different grammatical category. Indeed, when we take either (*u-* | *a-*) as a point of reference and compare bigrams with the OBJ markers, we again find similar biases for prefixes of the same category (Figure 3).

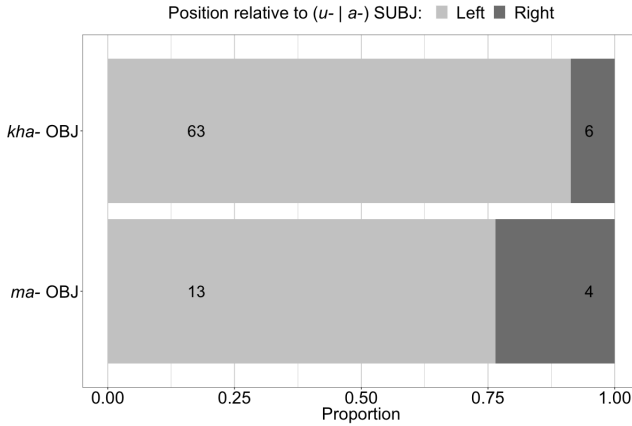


FIGURE 3. Bigram sequences where the reference prefix *u-* or *a-* combines with prefixes of the same grammatical category, OBJ.

The biases in bigram sequencing of {SUBJ, NEG} (Fig. 2) and {OBJ, SUBJ} (Fig. 3) suggest a probabilistic template OBJ > SUBJ > NEG, with agreement markers aligning by role rather than person. For other bigram types the token counts are rather low, but they remain consistent with this template. *kha-* 1NSG.P and *mai-* NEG are observed together in just seven tokens, but these all select the OBJ-NEG sequence. *na-* 3>2 and *mai-* NEG have twenty-five tokens, skewed 19 : 6 toward the (OBJ+SUBJ)-NEG sequence. For bigrams of *ma-* 1NSG.P with *mai-* NEG, the data on sequences are not sufficiently clear due to homophony or near-homophony of markers. There is also a notable scarcity of the latter bigrams, which turns out to be interesting in itself, as it suggests a form of probabilistic homophony avoidance (see Supplementary Material 1).

To test statistically whether same-category prefixes such as *u-*, *a-* do indeed have the same placement bias (paradigmatic alignment), and different category morphs such as *kha-*, *mai-* do indeed have different placement biases (featural coherence), we fit a logistic regression model for bigram sequences. However, when modeling the probabilities of a given prefix sequence at a given point in time, we need to control for persistence effects between discourse tokens that might skew these probabilities. We explain this in more detail before we develop our model.

**3.4. PERSISTENCE BETWEEN TOKENS.** It has been previously observed that prefix ordering in Chintang is subject to priming effects, whereby a prefix bigram is likely to repeat the same sequence as was used in a recently uttered verb (Bickel et al. 2007:64). However, it is possible that these effects are a special case of a larger persistence effect (Szmrecsanyi 2006), that is, general tendencies for consecutive discourse tokens to match for some variable. For the purpose of statistical modeling, this more general effect is a more stringent control since it reduces the chances of detecting category clustering as a spontaneous production effect, that is, it works more strongly against our hypothesis than a more narrowly defined priming effect.

There is indeed evidence for persistence effects in the corpus data. Bigrams with sufficient tokens to allow testing include SUBJ and NEG prefixes. Of the 445 tokens with this type of bigram, about half ( $N = 229$ ) have at least one preceding token observed in the same recording session. Figure 4 charts these tokens according to the distance between preceding and subsequent token, measured by counting transcribed utterance breaks.

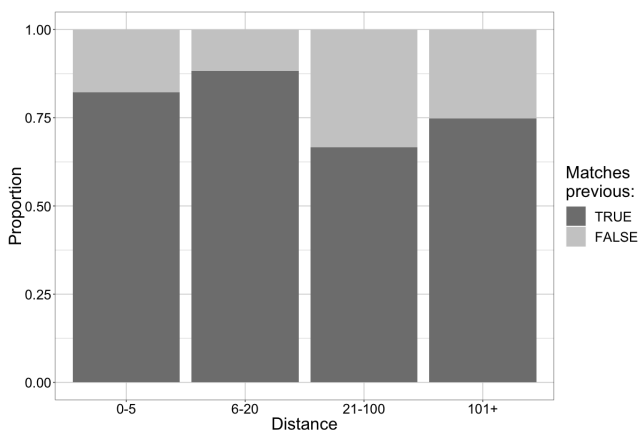


FIGURE 4. Persistence of {SUBJ, NEG} prefix bigrams by distance in annotation units (roughly corresponding to clauses).

Figure 4 shows that successive {SUBJ, NEG} bigrams have a very high chance of matching in prefix sequence. We can see that this matching is statistically significant by comparing it against the chance of any random pair of observations in the sample having matched {SUBJ, NEG} order. The chance of random pairs matching is 0.55,<sup>11</sup> but the rate of matching in discourse-consecutive pairs, 0.78, is significantly higher (binomial test: 95% CI = [0.72, 0.84], one-sided  $p < 0.001$ ). The persistence effect is especially pronounced for pairs of tokens separated by twenty utterances or fewer, though matching is also significant for longer-distance pairs. Consecutive tokens separated by over 100 utterances involve a passage of time that may be too great to be cognitively relevant (Szmrecsanyi 2006, Travis 2007).

**3.5. A STATISTICAL MODEL OF PROBABILISTIC CATEGORY CLUSTERING.** We focus on the placement of the SUBJ markers with respect to OBJ and NEG markers they cooccur with because unlike other prefix combinations, this provides a sufficiently large data set of  $N = 531$  bigrams. We test category clustering with a multilevel (‘mixed-effects’) logistic regression model, a standard tool for testing the effect of various variables on a binary response while controlling for others (Baayen 2008, Levshina 2015). In order to allow richer and more flexible model specification and evaluation, we use a Bayesian approach (Bürkner 2017), while providing a more traditional frequentist analysis in Supplementary Material 3 (which also contains further details on the Bayesian model). The response variable is the (log) odds of placing a SUBJ prefix to the left (vs. to the right) in a bigram.

<sup>11</sup> This figure is derived by enumerating all possible combinations of tokens from the full set, which contains 296 SUBJ-NEG and 149 NEG-SUBJ. This yields  $296 \times 149 = 44,104$  nonmatching pairs, from a total set of  $\binom{445}{2} = 98,790$  combinations, and 54,686 matching pairs.

Two predictor variables capture the effects of paradigmatic alignment and featural coherence, respectively (Table 4). Paradigmatic alignment predicts that the odds of leftward placement are the same for the two SUBJ markers. We model this by defining a variable PREFIX IDENTITY that estimates the effect that specific SUBJ prefix identity, either *u-* and *a-*, has on the prefix's probability of leftward placement. We compare each prefix against an intercept fixed at log odds = 0, that is, as deviations from a uniform 0.5 probability for each prefix to occur to either the left or the right in the bigram. Paradigmatic alignment predicts that these same-category affixes should have similar coefficient estimates, that is, share the same bias in left vs. right positioning. Featural coherence predicts that the odds of leftward placement of a SUBJ marker significantly depends on what other marker it cooccurs with, its CO-PREFIX. Given the probabilistic OBJ > SUBJ > NEG template proposed above, we specifically expect that SUBJ markers are less likely to be placed on the left in a bigram with an OBJ co-prefix than in a bigram with a NEG co-prefix. Whereas prefix identity levels are each compared against a zero intercept, for other predictor variables we compare one or more contrast levels against a reference level. We take NEG co-prefixes as the reference level for the co-prefix factor because left vs. right SUBJ placement is more equibaised in {SUBJ, NEG} than in {SUBJ, OBJ} bigrams. We arbitrarily select *u-* as the reference prefix identity for calculating estimates of other variables (this has no bearing on the results).

In order to control for persistence effects, we include a variable PERSISTENCE. We set the value 'no preceding token' as the reference level, and test the effect of a preceding same-categories bigram (separated by any number of annotation units) having SUBJ on either the left or the right. A preceding leftward-placed SUBJ marker is predicted to have a positive effect on leftward placement in the current bigram; a preceding rightward placement is predicted to have a negative effect. We furthermore model the effect of variation by individual lexical stems and speakers through random intercepts and random slopes. This ensures that any effects due to lexical choice or speaker habits are not erroneously attributed to the explanatory variables (Baayen 2008, Levshina 2015).<sup>12</sup> We choose a skeptical prior to favor the null hypothesis of no effect (see Supplementary Material 3).

PREDICTOR	TREATMENT LEVELS	REFERENCE LEVEL OR INTERCEPT	PREDICTED EFFECT OF TREATMENT
PREFIX IDENTITY	<i>u</i> - <sup>a</sup>	log odds = 0	positive
	<i>a</i> -		positive
CO-PREFIX	OBJ	NEG	negative
PERSISTENCE	Left	No preceding token	positive
	Right		negative

TABLE 4. Predictor variables used in the regression model of the (log) odds of placing a SUBJ marker to the left in a bigram with some co-prefix.

<sup>a</sup> *u-* is the reference level used for estimating the effect of the other variables.

The model's estimates support our hypothesis in all respects. Figure 5 shows the log odds coefficients estimated by the model. Whisker-lines represent credibility intervals (CI) in terms of the 95% highest density of posterior estimates. The SUBJ prefixes *u-* and *a-* have very similar estimates, each equally favoring leftward placement in a bigram,

<sup>12</sup> Chintang verb agreement varies across lexical valency classes (Schikowski et al. 2015). Because of this, we investigated whether {SUBJ, NEG} ordering varies across lexical stems. We found that for those lexical stems with sufficient token counts, the ordering was always quite close to the overall mean. This suggests that prefix ordering is largely independent of lexical stems, consistent with the findings of Ryan (2010) for Tagalog.

and excluding zero (no effect) from their credible estimates (*u-* 95% CI = [0.18, 1.60]; *a-* 95% CI = [0.40, 1.84]).<sup>13</sup> This supports paradigmatic alignment. Changing the co-prefix from NEG to OBJ has a clear negative effect; that is, the odds for leftward placement of SUBJ markers are much lower in bigrams with OBJ markers than in those with NEG markers (95% CI = [-5.17, -1.57]). This supports featural coherence in line with the OBJ > SUBJ > NEG template.

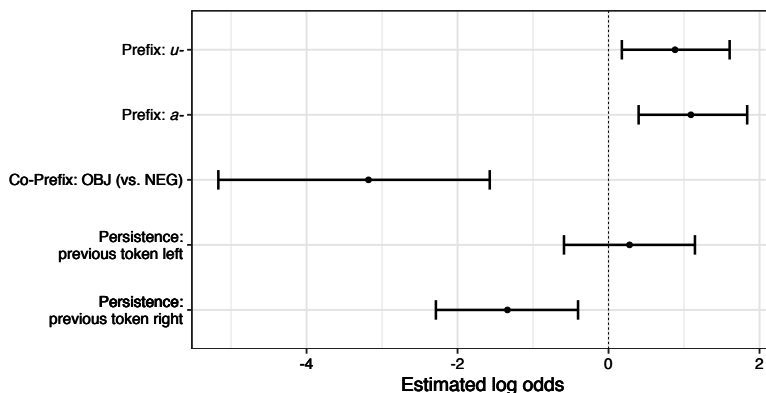


FIGURE 5. Coefficients, with 95% credibility intervals, estimated by Bayesian multilevel logistic regression model on the (log) odds of placing a SUBJ prefix on the left of a prefix bigram.

Importantly, these effects are present independently of any other variables. Persistence has a significant effect so that previous rightward placement of SUBJ markers decreases the odds for leftward placement (95% CI = [-2.29, -0.40]). However, previous leftward placement does not clearly increase over the baseline odds for a SUBJ prefix to be on the left (95% CI = [-0.59, 1.14]). The co-prefix effect remains independently strong, and there is no interaction of prefix identity with either co-prefix (95% CI = [-1.36, 2.40]) or persistence (95% CI = [-0.81, 1.55]).<sup>14</sup> Speaker and lexeme random effects account for a relatively small amount of variance compared to the main effect of the co-prefix.<sup>15</sup>

In summary, Chintang prefix bigram sequences exhibit clear biases to prefer one sequence over another. These biases conform to the principle of category clustering, exhibiting both paradigmatic alignment and featural coherence effects. Prefixes of the same category have similar placement biases, as expected from paradigmatic alignment. But prefixes of different categories have different biases, preserving featural coherence in a probabilistic fashion. Our model shows that these biases remain significant when we control for persistence effects and for random effects of speaker identity and lexical stem.

<sup>13</sup> Furthermore, assuming a model with an intercept produces posterior estimates close to zero, that is, no difference between the two (95% CI = [-0.07, 0.16]; 82% of estimates are between -0.1 and 0.1); see Supplementary Material 3.

<sup>14</sup> We were not able to test the interaction between co-prefix and persistence, since the co-prefix = OBJ group did not have any tokens with persistence = Left, which is unsurprising since rightward placement of SUBJ is dominant where co-prefix = OBJ, and {SUBJ, OBJ} bigrams are sparse enough in discourse that most tokens have no preceding token. This makes it very unlikely that the effect of the co-prefix is driven by persistence.

<sup>15</sup> The lexical stem has relatively low standard deviation both as a random intercept (median posterior estimate 0.41) and as a random slope (0.73). By contrast, individual speakers vary more widely, in standard deviation of both the random intercept (1.78) and the random slope (1.10).

While our model provides clear evidence for probabilistic clustering of Chintang prefixes, it remains to be seen whether this can be shown for other languages with free affix order. The Chintang corpus is one of the few large corpora available for languages with free affix placement, and the development of similar corpora for other such languages can be expected to yield further insights into the probabilistic structure of affixation.

**4. TYPOLOGICAL EVIDENCE FOR CATEGORY CLUSTERING.** We have seen above that Chintang prefixes exhibit probabilistic category clustering, showing that a widely assumed principle of morphology extends beyond fixed grammatical systems to those where different orders are equally grammatical. But we have also noted above that fixed-position systems sometimes exhibit nonclustering. If category clustering is a universal cognitive bias, we predict that when fixed-position systems evolve over time, it is more likely for them to comply with clustering than not. Concretely, when languages develop new agreement markers (e.g. by reanalyzing clitic pronouns or auxiliaries as affixes), we expect that markers of the same category, for example, subject markers, tend to be kept in the same position. When new markers develop in addition to already existing agreement marking, we expect that the earlier markers are likely to erode or to fuse with the new ones, so that the system does not end up with markers in different positions, violating clustering. Together, these developments should bias diachronic pathways to such an extent that synchronic systems tend to display category clustering beyond what one would expect by chance, after controlling for phylogenetic relations between languages.

We test this prediction against AUTOTYP data on grammatical markers (Bickel, Nichols, et al. 2017, file ‘Grammatical\_markers.csv’),<sup>16</sup> which codes various properties of individual grammatical markers that were collected for various purposes (surveys of agreement and case morphology, tense and plural marking, etc.).

**4.1. STUDY DESIGN.** AUTOTYP includes sufficiently detailed morphological placement data for agreement markers on verbs, though not for other types of grammatical markers (cf. Witzlack-Makarevich et al. 2016). We therefore test paradigmatic alignment by focusing on the extent to which agreement markers of the same category share the same morphological position, and test featural coherence by measuring the extent to which morphological positions host agreement markers of the same category. However, we are not able to test for alignment or coherence with respect to other types of affixes.

Testing the paradigmatic alignment effect (2a) requires a null model of random placement. This allows us to quantify the extent to which the observed alignments in the AUTOTYP data exceed chance and therefore support the idea of a universal alignment bias. To model random affix placement, we require a schema for possible verbal affix positions for each language tested. We do this based on known affix positions extracted from AUTOTYP, and our measure of alignment is relative to the number of known positions. In a verb with many attested positions, the alignment of same-category affixes is less probable, while in a verb with just two or three positions, alignment is more likely to occur by chance. The total number of affix positions can be inferred from agreement position indices given in the database; for example, a marker recorded in position 3 implies the existence of positions 1 and 2, even if we do not know what markers appear in those other positions. But this does not allow us to infer the existence of affix positions that are more peripheral than any agreement marker, and we therefore

<sup>16</sup> We used an advance copy of AUTOTYP release version 0.2, and our extracted paradigms are all listed in the appendix.

tend to underestimate the number of affix positions. Since nonrandom alignment is statistically more difficult to detect when there are fewer positions, this limitation makes our test design particularly conservative.

Testing the featural coherence effect (2b) requires an assessment of the extent to which positions contain markers of the same category. A language might comply with paradigmatic alignment by placing nearly all agreement markers in the same position, but if the language has a general preference for that position, other markers will also occur there and the position is no longer featurally coherent (cf. Table 1). As noted above, we do not have sufficient data on other categories for a full test, but it is possible to test featural coherence at least with regard to agreement marker categories. Below we assess the extent to which morphological positions differentiate between markers of the A (most agent-like) vs. the P (most patient-like) argument when a language has both types. We do this by measuring the statistical association of affix positions with categories.

**4.2. DATA OVERVIEW.** The full AUTOTYP grammatical markers data set contains 4,583 grammatical markers from 806 languages. The data set is not balanced for language family (for example, it contains large numbers of Algonquian and Kiranti languages), though we control for this in our analyses below. We focus specifically on the role categories of verbal agreement markers, as this mirrors the type of clustering found in our Chintang study, and because this is where AUTOTYP codes morphological positions most extensively.

AUTOTYP identifies role categories distinguishing between S (sole argument of one-place predicates),  $A_{tr}$  (most agent-like argument of two-place predicates),  $A_{ditr}$  (most agent-like argument of three-place predicates), P (least agent-like argument of two-place predicates), G (most location- or recipient-like argument of three-place predicates), and T (least location- or recipient-like argument of three-place predicates), cross-classified by lexical predicate classes (Bickel et al. 2014, Bickel, Nichols, et al. 2017, Witzlack-Makarevich et al. 2016). There are only a few agreement markers in the database for the  $A_{ditr}$ , G, and T roles, while  $A_{tr}$  and S are in most languages expressed by the same markers.<sup>17</sup> Where multiple exponence occurs, each exponent is a separate data point (e.g. if first-singular P is jointly marked by two affixes, both of these are registered as members of the P paradigm). In order to have a sufficiently large test set and at the same time to avoid duplicating counts when the same markers are involved, we extract only  $A_{tr}$  (henceforth simply A) and P markers. We furthermore exclude paradigms from Chintang (Bickel et al. 2007) and Bantawa (Doornenbal 2009) that are known to have free variation in affix placement. After these exclusions, we end up with data for 216 agreement marker paradigms, drawn from 136 different languages in forty-four different language families. Both A and P paradigms are present in eighty of the languages, while fifty-three languages have A only, and three have P only. ‘Paradigm’ here refers to a set of complementary affixes marking distinct values for A and P, and not the entire set of all inflectional forms for a lexeme, as in Stump 2001. The appendix lists the full set of paradigms, while the raw data and scripts used in this study can be downloaded from GitHub.<sup>18</sup>

Figure 6a illustrates the number of attested positions available on the verb in each of the 136 languages, while Figure 6b illustrates the number of positions occupied by each of the 216 agreement marker paradigms. Most languages in our data have one to three positions available for affix placement. Turning to paradigms, single-position align-

<sup>17</sup> This instantiates the well-known preference for accusative (S = A) alignment in verb agreement (Bickel 2011, Siewierska 2004).

<sup>18</sup> <https://github.com/jbmansfield/Category-Clustering>

ment is by far the most frequent pattern ( $N = 108$ ), with progressively fewer paradigms using more positions than this.

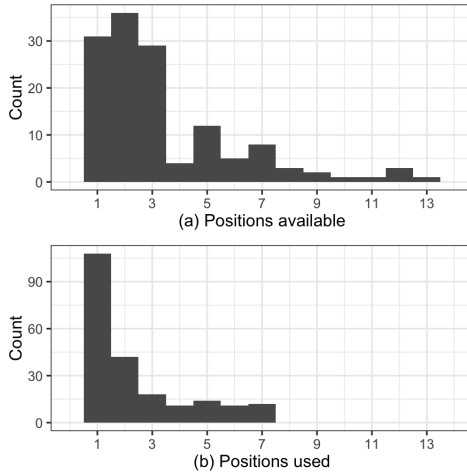


FIGURE 6. (a) Number of verbal affix positions available in 136 languages. (b) Number of affix positions used in 216 verbal agreement paradigms.

As expected, Fig. 6b shows that agreement paradigms generally use fewer positions than are available, with other available positions occupied by other inflectional categories. This suggests, on the one hand, that paradigmatic alignment effects are present in most of the languages sampled. On the other hand, we find single-position or ‘absolute’ alignment in only half of our paradigms, with various degrees of misalignment in the remainder. This shows the need for a nuanced test of whether the degree of alignment observed is beyond chance.

**4.3. A STATISTICAL MODEL OF PARADIGMATIC ALIGNMENT.** In order to probe the evidence for a paradigmatic alignment bias, we need to estimate the probability that a given alignment occurs against a null model under random, unbiased placement. For example, AUTOTYP records five A affixes for Reyesano (Pano-Tacanan; Guillaume 2009) and three affixal positions, two before and one after the stem ( $\Sigma$ ), as illustrated in Table 5. The Reyesano A affixes are not all allocated to the same position: four of them are in the prefix position  $\Sigma_{-2}$  and one in the suffix position  $\Sigma_{+1}$ . Intuitively, such an allocation suggests some degree of paradigmatic alignment, but this could be due to chance. To quantify effects beyond chance we need to first estimate the probability of observing a given allocation under a null model, and then rank the allocations by their degree of alignment. We take up these issues in turn.

	$\Sigma_{-2}$	$\Sigma_{-1}$	( $\Sigma$ )	$\Sigma_{+1}$
1SG	m-			
1PL	k-			
2SG	mi-			
2PL	mik-			
3				-ta

TABLE 5. Reyesano A affix allocations (Pano-Tacanan; Guillaume 2009).

To calculate the probability of a given allocation under a null model, we consider all logically possible allocations within the language, given the number of affixes and the

number of available positions.<sup>19</sup> We are indifferent to which particular positions are in fact selected—our interest lies only in the degree to which affixes are placed in the same position(s). In other words, we treat an allocation with four markers in  $\Sigma_{-2}$  and 1 in  $\Sigma_{+1}$  as showing the same degree of paradigmatic alignment as an allocation with one marker in  $\Sigma_{-2}$  and four in  $\Sigma_{+1}$ . Given this, the possible allocations are grouped according to the cardinality of the groupings over positions, that is, the mathematical partition of the paradigm (Hardy & Wright 2008:362ff.). For example, the five Reyesano A affixes can be partitioned into the available positions as the (unordered) sets  $\{5\}$ ,  $\{4, 1\}$ ,  $\{3, 2\}$ ,  $\{3, 1, 1\}$ , or  $\{2, 2, 1\}$ . As illustrated in Table 6, some partitions are produced in many different ways, while others are produced in only a few ways. Therefore, under the null model of random placement, some partitions are more probable than others. As a general rule, more highly aligned partitions (e.g.  $\{5\}$ ,  $\{4, 1\}$ ) are satisfied by fewer possible allocations, and are therefore less probable under random placement.

The formula for calculating the number of different allocations that produce each partition is described in Supplementary Material 2. In a nutshell, it involves calculating how many ways a set of affixes can be grouped to give a particular partition, and how many ways these groups can be distributed over the available positions. Given the number of possible allocations for a partition, the probability of that partition is its proportion of all possible allocations. For example, the Reyesano A  $\{4, 1\}$  partition accounts for 30/243 of all possible allocations, giving it a probability of 0.12.

In order to rank the partitions according to their degree of paradigmatic alignment within a given language, we draw on information entropy (Shannon 1948). Information entropy represents the degree to which a distribution of elements is nonuniform, that is, biased and predictable; it is calculated by summing the weighted log probabilities of each element. Lower entropy means a biased distribution, that is, more predictable outcomes, resulting either from a smaller set of elements, or from one element being much more probable than the others. If we treat affix allocations over positions as distributions of elements,  $\{5\}$  is the most biased distribution, with an entropy of  $H = 0$ , and this corresponds to full paradigmatic alignment. An allocation like  $\{3, 2\}$ , by contrast, is less biased and therefore has a higher entropy of  $H = 0.97$ ,<sup>20</sup> indicating considerably less alignment. While the absolute entropy values are not of interest for our purposes, they allow us to rank partitions according to their biases, that is, their degree of paradigmatic alignment. For example, the partition  $\{4, 1\}$  has a lower entropy ( $H = 0.72$ ) and therefore a higher degree of alignment than the partition  $\{3, 2\}$ .

The entropy-based ranking of partitions allows us to derive the cumulative probability of observing a given partition with a given degree of alignment, or a partition with any higher degree of alignment. For example, the observed Reyesano partition  $\{4, 1\}$  has a probability of 0.12 under the null model of random placement, but the cumulative probability of observing this much alignment or more is the sum of both  $\{4, 1\}$  and  $\{5\}$  probabilities, that is,  $0.12 + 0.01 = 0.13$ . Table 7 again shows the possible partitions of Reyesano A affixes, now with the figures for entropy, probability, and cumulative probability.

<sup>19</sup> Our model assumes that every affix can go in any position. This is not strictly true for pairs of affixes involved in multiple exponence, since these occur together and therefore by definition cannot go in the same position. Our model of possible distributions therefore includes a small number of clustered distributions that are not strictly plausible, and in doing so makes clustering appear more likely than it would otherwise. In summary, this slightly exaggerates the probability of the null model against which we provide evidence, making our test even more conservative.

<sup>20</sup>  $H = -\sum p \times \log_2 p$ , that is, in this case  $H = -\left(\frac{2}{5} \times \log_2 \left(\frac{2}{5}\right) + \frac{2}{5} \times \log_2 \left(\frac{2}{5}\right)\right) = 0.97$ .



PARTITION	EXAMPLE ALLOCATIONS (See also Supplement 2.)			# OF ALLOCS	PROB
	$\Sigma_{-2}$	$\Sigma_{-1}$	$\Sigma_{+1}$		
{5}	1SG, 1PL, 2SG, 2PL, 3	—	—	3	0.01
	—	1SG, 1PL, 2SG, 2PL, 3	—		
	—	—	1SG, 1PL, 2SG, 2PL, 3		
{4, 1}	1SG, 1PL, 2SG, 2PL	3	—	30	0.12
	1SG, 1PL, 2SG, 3	2PL	—		
	<i>etc. ...</i>	—	—		
	—	1SG, 1PL, 2SG, 2PL	3		
	—	1SG, 1PL, 2SG, 3	2PL		
	<i>etc. ...</i>	—	—		
	3	1SG, 1PL, 2SG, 2PL	—		
	2PL	1SG, 1PL, 2SG, 3	—		
{3, 2}	1SG, 1PL, 2SG	2PL, 3	—	60	0.25
	1SG, 1PL, 2PL	2SG, 3	—		
	1SG, 2SG, 2PL	1PL, 3	—		
	<i>etc. ...</i>	—	—		
	—	1SG, 1PL, 2SG	2PL, 3		
	—	1SG, 1PL, 2PL	2SG, 3		
	—	1SG, 2SG, 2PL	1PL, 3		
	<i>etc. ...</i>	—	—		
{3, 1, 1}	1SG, 1PL, 2SG	2PL	3	60	0.25
	1SG, 1PL, 2PL	2SG	3		
	1SG, 2SG, 2PL	1PL	3		
	<i>etc. ...</i>	—	—		
	1SG, 1PL, 2SG	3	2PL		
	1SG, 1PL, 2PL	3	2SG		
	1SG, 2SG, 2PL	3	1PL		
	<i>etc. ...</i>	—	—		
{2, 2, 1}	1SG, 1PL	2SG, 2PL	3	90	0.37
	1SG, 2SG	1PL, 2PL	3		
	1SG, 2PL	1PL, 2SG	3		
	1SG, 1PL	2SG, 3	2PL		
	1SG, 2SG	1PL, 3	2PL		
	<i>etc. ...</i>	—	—		
	<i>etc. ...</i>	—	—		
TOTAL				243	1.00

TABLE 6. Possible allocations of Reyesano A markers over available positions under a null model of random placement.

	ENTROPY	PROB	CUMULATIVE PROB
{5}	0	0.01	0.01
{4, 1}	0.72	0.12	0.13
{3, 2}	0.97	0.25	0.38
{3, 1, 1}	1.37	0.25	0.63
{2, 2, 1}	1.52	0.37	1.00

TABLE 7. Reyesano A affixes partitions, entropy, and probability.

We use the cumulative probability as a paradigmatic alignment index for each paradigm, converted into a positive value on the scale 0 to 1 by subtracting the cumulative probability from 1. Thus the higher the index, the greater the degree of paradigmatic alignment, relative to all possible allocations under random placement. Reyesano A markers have a fairly high paradigmatic alignment index of 0.87: the observed alloca-

tion of {4, 1} or one with even more alignment is unlikely (with probability  $1 - 0.87 = 0.13$ ) to occur by random placement.

In calculating paradigmatic alignment indices, we exclude paradigms from languages that have only a single known affix position on the verb. In these languages markers align trivially even in the absence of any category clustering bias. We therefore exclude them from our test, and this reduces the data set to 180 paradigms, 105 languages, and thirty-eight language families. Excluding single-position languages is again a conservative measure that goes against our hypothesis, since this reduces the overall degree of alignment in the data. The full list of paradigms is provided in the appendix, including both those with only a single known position and those with multiple positions, the latter listed with paradigmatic alignment index scores.

Figure 7 shows the distribution of the paradigmatic alignment index for A and P markers. In both categories, there is an apparent bias toward high values, and therefore high degrees of paradigmatic alignment in most paradigms. A markers also show a group of paradigms with very low paradigmatic alignment indices (i.e. close to zero), an observation to which we return after statistically modeling the distributions.

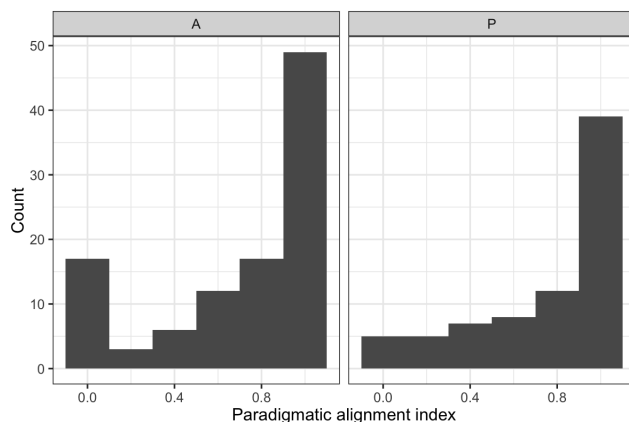


FIGURE 7. Degrees of paradigmatic alignment for A and P roles.

To test whether the biases in Fig. 7 reflect a statistical bias, we set up a multilevel mixture model on the paradigmatic alignment index, estimating at the same time the alignment values between 0 and 1 with a beta regression and the probabilities of 0 alignment with a logistic regression. While the logistic component is a common choice in language science, beta regression is less common. It is designed for cases where the outcome is constrained to the unit interval between but excluding 0 and 1, and it follows a beta distribution, but is not meaningfully transformable into binary odds or probabilities (Cribari-Neto & Zeileis 2010, Ferrari & Cribari-Neto 2004). In all other regards the model follows the same logic as any other regression. As in the model for Chintang prefix placement above, we capture the main effects of interest by comparing each marker category, A and P, as deviations from equal probability (i.e. a 0.5 mean, corresponding to a logit = 0). We control for phylogenetic autocorrelation by including language family as a random intercept.<sup>21</sup>

<sup>21</sup> We also considered testing geographical area as an additional random factor, but we refrained from doing so because of the high degree of collinearity between region and language family.

We fitted the mixture model in a Bayesian framework with a skeptical prior (Supplementary Material 3). The beta component suggests that both A and P categories have estimates at the upper end of the index (both A and P have median posterior estimates of 0.86), and they exclude neutral 0.5 values from their 95% credibility intervals by a fair margin (95% CIs A = [0.81, 0.91], P = [0.80, 0.92]). The logistic component of the mixture model furthermore reveals that the estimated probabilities for zero values are exceedingly small (median for A = 0.037 and for P = 0.003). These estimates are notably lower than what the marginal counts suggest in Fig. 7. This is due to the fact that the model controls for the historical relationships between languages in the random effects, while the figure overcounts data from related families. Indeed, we note a high standard deviation estimate of the random intercept both for values between 0 and 1 (95% CI = [0.63, 0.76]) and for the probability of 0 values (95% CI = [0.76, 1]). Taken together, these results suggest strong biases toward paradigmatic alignment in both A and P categories (see Supplementary Material 3 for further details on the model).

**4.4. NONALIGNED PARADIGMS.** While our test confirms a general bias toward paradigmatic alignment, we also note that before historical relationships are controlled for, a substantial number of paradigms exhibit nonclustering. These are paradigms in which placement is dispersed relatively evenly across all known positions, approaching maximum possible entropy. There are a total of thirty-seven (out of 180) paradigms that have alignment indices below 0.5, and inspection of these nonaligned or ‘dispersed’ paradigms reveals that they fall into three groups.

The first group of dispersed paradigms ( $N = 14$ ) includes those for which only two positions have been identified, and A markers are evenly divided between these two. Several of these are in Berber languages, reflecting the split prefix/suffix marking shown for Tamazight Berber above (§2.3), which has a {4, 3} partition in two known positions. These paradigms approach maximum dispersion because the affixes are evenly distributed among all known positions, though in absolute terms they do not involve very extensive dispersion.

The second group ( $N = 17$ ) involves agreement roles being evenly divided over a large number of positions, mostly in Algonquian and Kiranti (Sino-Tibetan). Many of the agreement-marking affixes in these languages seem to be aligned by person rather than role. For example, Cheyenne has thirteen A affixes, spread over all seven known positions in the partition {3, 3, 2, 2, 1, 1, 1}, and fifteen P affixes with a similar dispersion {3, 3, 3, 2, 2, 1, 1}. These are close to maximum dispersion by role. Some of the relevant markers might show stronger alignment by person instead (Goddard 2000), reflecting the often posited tendency of Algonquian and Kiranti languages to show person-driven alignment (DeLancey 1981, Ebert 1987, Hockett 1966, Nichols 1992). However, as shown by Witzlack-Makarevich et al. (2016), the evidence for person-driven alignment is in fact quite weak in these languages, from both a synchronic and a diachronic perspective. Consistent with this, we find that several Algonquian and Kiranti languages do have paradigmatic alignment in terms of role (see the appendix).

The third group ( $N = 6$ ) are Kiranti agreement paradigms with extensive multiple exponence (Harris 2017), where the same category is simultaneously expressed in several positions. For example, Yakkha has seven A markers spread over seven positions (with a total of thirteen positions attested), that is, the maximum-entropy partition {1, 1, 1, 1, 1, 1, 1}. Inspection of the Yakkha verb template shows that almost every inflectional affix in the language has its own position, because A, P, and TAM features are generally encoded in a distributed fashion over sequences of suffixes (Schakow 2015:207). For example, a transitive verb with 1DU.EXCL > 3NSG spreads person/number markers over

four affixes, which must therefore each have their own position (21). Highly distributed affix systems of this type may be simultaneously misaligned for all features.

(21) *tund-aŋ-c-uŋ-ci-ŋ*(=ha)<sup>22</sup>

understand-PST-DU-3.P-3NSG.P-EXCL=NMLZ.NSG

‘We (dual, excl.) understood them.’

(Schakow 2015:219)

A likely diachronic source of such patterns is repeated fusion of auxiliaries, each with their own agreement markers. When this happens without erosion of earlier markers, category clustering is very limited and the resulting system is an idiosyncratic affix template.

**4.5. A STATISTICAL MODEL OF FEATURAL COHERENCE.** As mentioned above, since the AUTOTYP data set contains only positional information on verbal agreement markers, the only opportunity we have to test for featural coherence is in languages for which the verb hosts agreement for multiple roles. We therefore focus on those languages that have both A and P agreement markers, testing whether these are aligned differently. For example, Mursi (Surmic) has a high paradigmatic alignment index for both A and P markers, and furthermore, these arguments are aligned quite differently (Table 8a). By contrast, in Teso (Nilotic), both A and P again have high paradigmatic alignment, but in this case they both align in  $\Sigma_{-1}$ , and therefore do not exhibit coherence (Table 8b).

	$\Sigma_{-1}$	$\Sigma_{+1}$	$\Sigma_{+2}$
A	1	0	5
P	0	4	0

TABLE 8a. Mursi: Paradigmatic alignment (shaded) and featural coherence.

	$\Sigma_{-2}$	$\Sigma_{-1}$	$\Sigma_{+1}$
A	0	5	1
P	1	3	0

TABLE 8b. Teso: Paradigmatic alignment (shaded) but NOT featural coherence.

We capture featural coherence by measuring the statistical association between markers and cells, using Cramér’s  $V$  corrected for biases induced by small samples in large tables (Bergsma 2013). This statistic assesses the extent to which counts in cells deviate from what is expected under a null model of no association (balanced distribution), corrected for the number of cells in a table. The statistic ranges from 0 to 1, with higher figures indicating a stronger association, in our case indicating that positioning is associated with semantic role. For example, Mursi has  $V = 0.94$ , while Teso has  $V = 0.00$ , reflecting the fact that A vs. P allocations are much less differentiated across positions in Teso. When we calculate  $V$  for eighty-two languages with multiple positions and both A and P markers,<sup>23</sup> we find that almost all languages measure toward the extremes of the scale, with just over half of the languages ( $N = 45$ ) showing high featural coherence ( $V \geq 0.5$ ), while the remainder have low measures of coherence (Figure 8). All featural coherence measurements are listed alongside alignment indices in the appendix.

Figure 8 shows similar numbers of languages with high and low featural coherence, suggesting that our sample may not have a bias toward coherence as was found for paradigmatic alignment. However, closer inspection of the data reveals that noncoherence for A vs. P categories is limited to two language families, Algonquian and Kiranti,

<sup>22</sup> A morphophonological process copies the *-ŋ* EXCL suffix regressively into other coda positions. Schakow annotates this as a series of extra affixes, but we interpret it here as separate from affix placement.

<sup>23</sup> This includes seventy-eight of the eighty languages with A and P paradigms mentioned above (the excluded two being single-position languages), plus four others in which there is both A and P agreement, but one or both categories have just a single marker and were therefore not treated as paradigms.

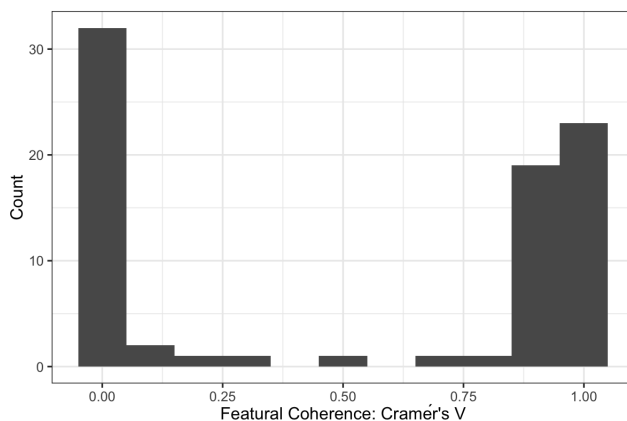


FIGURE 8. Featural coherence of A and P paradigms as measured by bias-corrected Cramér's  $V$ .

which also account for many of the nonaligned paradigms observed above. Noncohering languages include some in which the paradigms are also nonaligned, such as Blackfoot, Cheyenne (Algonquian), Athpare, and Yakkha (Kiranti). But other noncohering languages do have paradigmatic alignment, such as Arapaho, Plains Cree (Algonquian), Dumí, and Wambule (Kiranti). The latter group tend to lack coherence because A and P align to the same positions, that is, positional competition of the type exemplified for Wópanâak (Algonquian) in 10.

Figure 9 shows featural coherence measures with Algonquian and Kiranti separated from all other language families. As this figure shows, there does in fact appear to be a coherence bias in most families, but it is altogether absent in Algonquian and Kiranti.

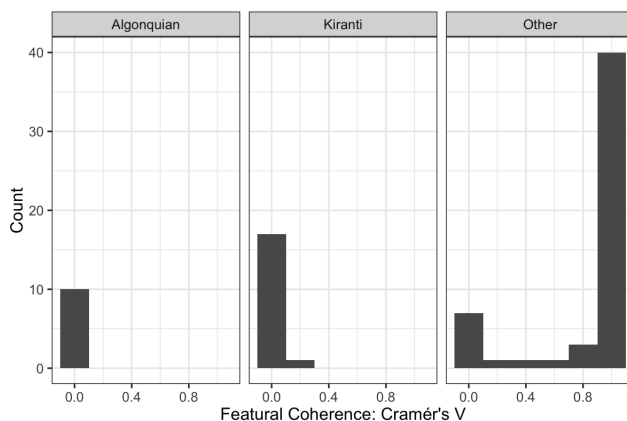


FIGURE 9. Featural coherence, with Algonquian and Kiranti separated from all other language families.

We test for a bias toward featural coherence by again using a multilevel mixture model, this time with Cramér's  $V$  as the response variable. The intercept is the coefficient of interest, with an intercept above 0.5 indicating a bias toward featural coherence. We have no predictor variables, but a random effect of language family to control for the high degree of variance shown in Fig. 9. The result supports our hypothesis. The beta component of the mixture model reveals (on the inverse logit, i.e. response scale) a median posterior estimate of  $V = 0.77$  (95% CI = [0.64, 0.89]). The logistic component

reveals that complete coherence ( $V = 1$ ) is much more likely (95% CI = [0.98, 1]) than incoherence ( $V = 0$ ). This suggests that the apparent high count of 0s in Fig. 8 is an artefact of historical dependencies between languages of the same family, and these are captured by the model through high random-effect estimates (see Supplementary Material 3 for details).

In summary, once historical relationships are controlled for, our typological data show strong positive biases toward both paradigmatic alignment and featural coherence. In most language families both of these biases are present. But two language families, Algonquian and Kiranti, have a mixture of aligned and nonaligned agreement paradigms, and none at all with featural coherence. Berber languages do not show paradigmatic alignment, though they were not relevant to our featural coherence test since they agree only for A participants.

**5. DISCUSSION AND OUTLOOK.** In our first study, we showed that Chintang prefixes exhibit a probabilistic bias toward category clustering. Although there are no grammatical rules determining prefix placement in this language, our corpus data suggest that in naturalistic language production, Chintang speakers are biased toward placing exponents of the same category in the same position (tending toward paradigmatic alignment) and different categories in different positions (tending toward featural coherence).

In our second study, we found a global preference for both A and P agreement markers to align in paradigmatic positions, rather than being scattered across positions. We also showed that at least for A and P agreement markers, languages tend to also comply with featural coherence, having specific positions for A and P each. However, some language families escape the general preference: Algonquian and Kiranti have many paradigms that defy both paradigmatic alignment and featural coherence; Berber has many paradigms that defy paradigmatic alignment (and we do not have enough Berber P paradigms to evaluate featural coherence).

Especially in the case of Algonquian and Kiranti, some of the deviations from the global trend potentially reflect clustering according to person instead of role.<sup>24</sup> However, further research is needed to establish the extent of this effect, because the evidence of person-based agreement morphology in these languages is considerably weaker than is sometimes claimed (Witzlack-Makarevich et al. 2016) and because highly dispersed exponence seems to be just as important (as in the Yakkha example of multiple exponence of agreement markers). At any rate, it remains a striking observation that in one case where a Kiranti language does not rigidly regulate affix placement, that is, in Chintang, there is again a bias toward clustering. This observation might suggest that Kiranti agreement systems are in a transitional phase of historical development, and that over the long run, deviations will be smoothed out by the same bias that lets Chintang speakers already cluster their prefixes at present. To resolve these possibilities, future research is needed, for example with artificial language learning experiments (cf. Culbertson et al. 2012) or morphological priming experiments (cf. Duñabeitia et al. 2009, Gagné et al. 2009).

These deviations notwithstanding, our global results confirm the assumption of category clustering as a default in morphological theory, at least with regard to agreement morphology. As such, they provide quantification methods for representing the clustering bias in formal models, for example, as a prior in probabilistic models or as a weighted term in symbolic approaches (e.g. Crysman & Bonami 2016, Sagot & Walther 2011). Furthermore, category clustering has implications for the debate about whether mor-

<sup>24</sup> We thank two referees for drawing our attention to this.

phology is separate from syntax. The principle of category clustering is shared with category-based phrase structure, and therefore the prevalence of clustering in affix placement may be adduced in support of a syntactic approach to morphology. Conversely, nonclustering presents a form of idiosyncratic affix placement that suggests morphological autonomy from category-based syntax.

From the categorial perspective that drives these debates, our findings are ambiguous: while our results on Chintang can be taken as evidence for a model of morphology that is similar to category-driven syntax, our typological study identifies both a global trend toward syntax-like clustering and a few recalcitrant deviations.

The ambiguity can be resolved if we follow Rice (2011:193) in concluding that ‘no single principle is able to account for all facets of [affix] ordering either between languages or within a language’. This supports the idea that category clustering, and indeed the similarity of morphology and syntax, is not a universal constant in the architecture of grammar but a typological variable. From this perspective, it is expected that some languages, or even some language families, deviate from category clustering. At the same time, however, one would not expect the variable to evolve completely at random, picking any value with equal probability. Instead, as is often the case in typological findings, its distribution is shaped by an underlying probabilistic principle, that is, category clustering, and so compliance with the principle is far more common than deviations.

If further studies can replicate the category clustering bias for other types of markers and for more languages, the question arises of what might cause the bias. One possible answer is predictability in language processing and acquisition. Processing requires fast categorization and prediction of various units within and between words. A key effect of category clustering is that exponents of the same category recur in the same morphological environment over and over again. This makes category identity more predictable, allowing the hearer to guess the category based on contextual cues before even hearing the relevant morpheme. This reasoning is supported by the persistence effect we find in the freely ordered prefixes of Chintang. Instead of changing the order of prefixes on a random basis within a conversation, speakers prefer to use the same ordering as uttered previously. This might increase predictability and hence processing speed.

Furthermore, learning inflected verb forms would be a much more difficult task if categories and the placement of their markers were hard to predict. Category predictions might help the child to categorize without yet knowing each exponent of the category. In corpus studies of child-directed speech it has been shown that the recurrent order of elements can indeed help in categorization. Frequently recurring ‘frames’ of surrounding elements may potentially help the child to identify the class of the middle element, that is, help in categorizing this element. This is what is known in acquisition studies as the ‘frequent frames’ effect (Chemla et al. 2009, Mintz 2002, 2003, Mintz et al. 2014), which has been shown to be a consistent property of the distribution of affixes in typologically maximally diverse languages, including Chintang (Moran et al. 2018).

If this explanation is on the right track, however, it is again puzzling that some languages seem to defy category clustering to a considerable extent. To resolve this puzzle, future research needs to go beyond agreement markers and assess whether these languages show clustering in other categories. Also, if a language violates category clustering, there might be strategies that compensate for the loss in efficiency for learning and processing. One such strategy might be precisely one of the patterns that leads to deeper violations of clustering in the first place: multiple, dispersed exponence. This is a prime means of establishing redundancy and it may have a beneficial effect for learning and processing. From this perspective, future research might profitably move be-

yond category clustering per se and instead assess directly to what extent the demands of learning and processing shape how languages order their affixes, with clustering and dispersion as different solutions to the same problem.

APPENDIX: LIST OF LANGUAGES AND PARADIGMS EXTRACTED FROM AUTOTYP<sup>25</sup>

Languages in *italics* have just one known affix position and are therefore excluded from the paradigmatic alignment bias calculation in §4.3.

LANGUAGE	STOCK	LANG ID	CAT	POS	POS	PARTITION	ALIGN	FEAT
				AVL	USED		INDEX	COH
Acehnese	Austronesian	9	A	3	1	{5}	0.99	1.00
			P	3	1	{5}	0.99	1.00
Ainu	Ainu	12	A	3	1	{4}	0.96	0.47
			P	3	2	{2, 2}	0.44	0.47
<i>Alaba-K'abeena</i>	Cushitic	3018	A	1	1	{6}	—	—
<i>Amanab</i>	Border	480	A	1	1	{2}	—	—
Amharic	Semitic	21	A	5	3	{6, 5, 2}	0.99	0.94
			P	5	1	{6}	1.00	0.94
Amuesha	Arawakan	885	A	3	2	{4, 1}	0.86	0.94
			P	3	1	{6}	1.00	0.94
Anamuxra	Madang	1645	A	4	1	{9}	1.00	1.00
			P	4	1	{9}	1.00	1.00
Anêm	West New Britain	22	A	2	1	{6}	0.97	1.00
			P	2	1	{7}	0.98	1.00
Arabic (Egyptian)	Semitic	642	A	3	3	{6, 5, 2}	0.53	—
Araki	Austronesian	871	A	2	1	{7}	0.98	1.00
			P	2	1	{1}	—	1.00
Arapaho	Algic	923	A	5	4	{8, 3, 2, 1}	0.98	0.00
			P	5	4	{8, 4, 2, 1}	0.99	0.00
<i>Armenian (Eastern)</i>	Indo-European	25	A	1	1	{6}	—	—
Asmat	Macro-Ok	26	A	2	1	{5}	0.94	1.00
			P	2	1	{2}	0.50	1.00
Atakapa	Atakapa	27	A	2	1	{4}	0.88	1.00
			P	2	1	{6}	0.97	1.00
Athpare	Sino-Tibetan	908	A	10	6	{2, 2, 2, 1, 1, 1}	0.35	0.00
			P	10	7	{3, 1, 1, 1, 1, 1, 1}	0.30	0.00
Atikamekw	Algic	2551	A	7	6	{3, 2, 2, 2, 2, 1}	0.23	0.00
			P	7	7	{3, 2, 2, 2, 2, 1, 1}	0.01	0.00
Awtuw	Sepik	28	A	9	2	{1, 1}	0.00	—
Baale	Surmic	1791	A	3	2	{6, 1}	0.98	—
Bahing	Sino-Tibetan	3007	A	6	4	{9, 1, 1, 1}	1.00	0.20
			P	6	5	{5, 2, 2, 1, 1}	0.76	0.20
<i>Bariai</i>	Austronesian	2982	A	1	1	{5}	—	—
Baure	Arawakan	1063	A	2	1	{6}	0.97	1.00
			P	2	1	{2}	0.50	1.00
Belhare	Sino-Tibetan	35	A	12	6	{2, 2, 2, 1, 1, 1}	0.51	0.00
		35	P	12	7	{2, 2, 1, 1, 1, 1, 1}	0.15	0.00
Berber (Figuig)	Berber	750	A	2	2	{4, 3}	0.00	—
Berber (Kabyle)	Berber	2882	A	3	3	{4, 3, 1}	0.45	0.80
			P	3	1	{6}	1.00	0.80
<i>Biak</i>	Austronesian	1014	A	1	1	{3}	—	—
Binandere	Greater Binanderean	1010	A	2	1	{3}	0.75	—
Blackfoot	Algic	1036	A	7	6	{2, 2, 2, 1, 1, 1}	0.06	0.00
			P	7	6	{2, 2, 2, 2, 1, 1}	0.10	0.00

(Appendix table continues)

<sup>25</sup> For an electronic version of this table, see <https://github.com/jbmansfield/Category-Clustering>.



LANGUAGE	STOCK	LANG ID	CAT	POS	POS	PARTITION	ALIGN	FEAT
				AVL	USED		INDEX	COH
<i>Bororo</i>	Macro-Ge	648	A	1	1	{6}	—	—
			P	1	1	{6}	—	—
Brahui	Dravidian	518	A	2	2	{6, 5}	0.00	—
Bulgarian	Indo-European	678	A	2	1	{6}	0.97	1.00
			P	2	1	{1}	—	1.00
Cakchiquel	Mayan	1155	P	2	1	{5}	0.94	—
Camling	Sino-Tibetan	2360	A	6	5	{2, 2, 1, 1, 1}	0.05	0.00
			P	6	6	{4, 1, 1, 1, 1, 1}	0.25	0.00
Chai	Surmic	1413	A	5	3	{4, 3, 1}	0.92	0.88
			P	5	1	{3}	0.96	0.88
Cheyenne	Algic	1142	A	7	7	{3, 3, 2, 2, 1, 1, 1}	0.08	0.00
			P	7	7	{3, 3, 3, 2, 2, 1, 1}	0.07	0.00
Choctaw	Muskogean	54	A	4	4	{4, 4, 3, 1}	0.41	0.10
			P	4	2	{4, 4}	0.98	0.10
Chontal Maya	Mayan	1136	A	5	3	{3, 3, 1}	0.83	0.29
Chortí	Mayan	1105	A	5	2	{3, 2}	0.90	0.29
			P	3	2	{5, 1}	0.95	0.94
Chuvash	Turkic	57	A	3	1	{5}	0.99	0.94
			A	2	2	{6, 5}	0.00	—
Cora	Uto-Aztecan	688	A	2	1	{5}	0.94	1.00
			P	2	1	{6}	0.97	1.00
Cree (Plains)	Algic	59	A	8	5	{5, 3, 2, 2, 1}	0.95	0.00
			P	8	4	{5, 4, 2, 2}	0.99	0.00
<i>Dagur</i>	Mongolic	1416	A	1	1	{6}	—	—
<i>Darmita</i>	Sino-Tibetan	1388	A	1	1	{3}	—	—
<i>Dogon (Ben Tey)</i>	Dogon	3092	A	1	1	{5}	—	—
<i>Dogon (Najamba)</i>	Dogon	3093	A	1	1	{5}	—	—
Dogon (Nanga)	Dogon	3096	A	2	1	{5}	0.94	—
Dumi	Sino-Tibetan	1439	A	5	3	{6, 2, 1}	0.99	0.00
			P	5	3	{6, 2, 1}	0.99	0.00
Emerillon	Tupian	3068	A	2	1	{6}	0.97	0.00
			P	2	2	{5, 1}	0.78	0.00
<i>French (colloquial)</i>	Indo-European	79	A	1	1	{2}	—	—
Ghomara	Berber	3039	A	2	2	{4, 3}	0.00	—
<i>Golin</i>	Chimbu-Wahgi	1578	A	1	1	{6}	—	—
Guaraní (Mbyá)	Tupian	3131	A	2	1	{6}	0.97	0.00
			P	2	2	{6, 1}	0.88	0.00
Gurage (Sebat Bet)	Semitic	3044	A	5	3	{6, 4, 3}	0.99	0.94
			P	5	1	{6}	1.00	0.94
<i>Hatam</i>	Hatam	645	A	1	1	{6}	—	—
Hayu	Sino-Tibetan	632	A	5	4	{3, 2, 1, 1}	0.38	0.00
			P	5	5	{3, 2, 2, 1, 1}	0.06	0.00
<i>Hebrew (Modern)</i>	Semitic	583	A	1	1	{4}	—	—
Hua	Eastern Highlands	103	A	3	2	{3, 3}	0.74	0.95
			P	3	1	{6}	1.00	0.95
Ik	Kuliak	111	A	2	1	{6}	0.97	—
Ineseño	Chumashan	113	A	3	2	{3, 2}	0.62	—
Itzaj	Mayan	1660	A	7	2	{4, 1}	0.99	0.93
			P	7	2	{6, 6}	1.00	0.93
Iyo	Finisterre-Huon	3062	A	5	1	{5}	1.00	0.93
			P	5	2	{3, 1}	0.86	0.93
Jacalteco	Mayan	460	A	2	1	{4}	0.88	1.00
			P	2	1	{4}	0.88	1.00
Jero	Sino-Tibetan	2998	A	3	3	{5, 3, 1}	0.68	0.00
			P	3	3	{2, 1, 1}	0.00	0.00
Juang	Austroasiatic	1691	A	3	3	{5, 2, 1}	0.70	0.67
			P	3	1	{6}	1.00	0.67

(Appendix table continues)

LANGUAGE	STOCK	LANG ID	CAT	POS	POS	PARTITION	ALIGN	FEAT
				AVL	USED		INDEX	COH
<i>Kamaiurá</i>	Tupian	1704	A	1	1	{9}	—	—
			P	1	1	{8}	—	—
<i>Karajá</i>	Macro-Ge	2951	A	1	1	{2}	—	—
Keresan (Laguna)	Keresan	2958	A	7	1	{2}	0.86	1.00
Khakas	Turkic	1763	A	4	1	{5}	1.00	—
			P	3	1	{8}	1.00	1.00
Khanty	Uralic	681	A	3	1	{3}	0.89	1.00
			P	3	1	{10}	1.00	—
Kharia	Austroasiatic	1750	A	2	2	{2, 2}	0.00	—
Koegu	Surmic	2772	A	3	3	{9, 4, 2}	0.90	0.00
Kõic	Sino-Tibetan	2956	P	3	3	{5, 4, 1}	0.69	0.00
			A	4	4	{5, 3, 1, 1}	0.69	0.00
Koyi	Sino-Tibetan	2980	P	4	3	{6, 5, 1}	0.98	0.00
			A	5	4	{2, 2, 1, 1}	0.12	0.00
Kulung	Sino-Tibetan	1775	P	5	4	{4, 4, 1, 1}	0.87	0.00
			A	1	1	{6}	—	—
<i>Latvian</i>	Indo-European	549	A	12	7	{2, 2, 1, 1, 1, 1, 1}	0.15	0.00
Limbu	Sino-Tibetan	674	P	12	7	{4, 1, 1, 1, 1, 1, 1}	0.67	0.00
			A	1	1	{4}	—	—
<i>Lithuanian</i>	Indo-European	1890	A	7	4	{3, 2, 2, 1}	0.73	0.00
Lohorong	Sino-Tibetan	3010	P	7	6	{3, 3, 2, 1, 1, 1}	0.37	0.00
			A	2	2	{6, 1}	0.88	0.00
Maa	Nilotic	167	P	2	1	{2}	0.50	0.00
			A	1	1	{6}	—	—
<i>Majang</i>	Surmic	2063	P	1	1	{9}	—	—
<i>Manambu</i>	Sepik	2028	A	11	7	{3, 3, 2, 2, 2, 1, 1}	0.72	0.00
Menomini	Algic	1973	P	11	7	{3, 2, 2, 2, 1, 1, 1}	0.53	0.00
			A	2	1	{7}	0.98	1.00
Menya	Angan	1954	P	2	1	{7}	0.98	1.00
			A	5	5	{7, 3, 1, 1, 1}	0.92	0.00
Micmac	Algic	2001	P	5	4	{5, 3, 3, 1}	0.85	0.00
			A	1	1	{4}	—	—
<i>Mixtec</i> ( <i>Chalcatongo</i> )	Otomanguean	186	A	2	2	{8, 7}	0.00	—
Moghol	Mongolic	2029	A	3	2	{4, 1}	0.86	0.91
Mugil	Madang	2031	P	3	1	{3}	0.89	0.91
			A	8	5	{3, 2, 2, 2, 1}	0.68	0.00
Munsee	Algic	2668	P	8	5	{3, 3, 3, 2, 2}	0.86	0.00
			A	6	3	{6, 3, 1}	1.00	0.91
Murle	Surmic	559	P	6	1	{4}	1.00	0.91
			A	3	2	{5, 1}	0.95	0.94
Mursi	Surmic	2098	P	3	1	{4}	0.96	0.94
			A	3	2	{3, 1}	0.67	0.94
Nahuatl (Sierra de Zacapoaxtla)	Uto-Aztecan	956	P	3	1	{6}	1.00	0.94
			A	3	2	{3, 1}	0.67	0.94
Nahuatl (Tetelcingo)	Uto-Aztecan	572	P	3	1	{6}	1.00	0.94
			A	1	1	{5}	—	—
<i>Nanai</i>	Tungusic	201	A	3	2	{5, 4}	0.91	0.95
Nandi	Nilotic	299	P	3	1	{4}	0.96	0.95
			A	1	1	{6}	—	—
<i>Nepali (Eastern)</i>	Indo-European	3117	A	2	1	{9}	1.00	1.00
Nganasan	Uralic	2172	P	2	1	{2}	0.50	1.00
			A	7	1	{4}	1.00	—
Nubian (Kunuz)	Nubian	1348	A	7	5	{4, 3, 2, 2, 2}	0.78	0.00
Ojibwa (Eastern)	Algic	2244	P	7	5	{4, 3, 2, 2, 2}	0.78	0.00
			A	2	1	{2}	0.50	1.00
Oksapmin	Macro-Ok	322	P	2	1	{2}	0.50	1.00

(Appendix table continues)

LANGUAGE	STOCK	LANG ID	CAT	POS	POS	PARTITION	ALIGN	FEAT
				AVL	USED		INDEX	COH
Old Thulung (Mukli)	Sino-Tibetan	2999	A	6	5	{4, 1, 1, 1, 1}	0.53	0.00
			P	6	5	{3, 3, 1, 1, 1}	0.53	0.00
Olo	Torricelli	2251	A	3	1	{7}	1.00	0.95
			P	3	2	{4, 1}	0.86	0.95
Passamaquoddy	Algic	563	A	12	6	{3, 3, 3, 2, 1, 1}	0.95	0.00
			P	12	6	{3, 3, 3, 2, 1, 1}	0.95	0.00
Persian	Indo-European	456	A	2	2	{6, 5}	0.00	—
Pipil	Uto-Aztecan	332	A	3	2	{3, 1}	0.67	0.94
			P	3	1	{6}	1.00	0.94
<i>Provençal</i>	Indo-European	2335	A	1	1	{5}	—	—
Puma	Sino-Tibetan	2863	A	8	5	{4, 3, 1, 1, 1}	0.90	0.09
			P	8	6	{4, 2, 1, 1, 1, 1}	0.61	0.09
Quechua (Imbabura)	Quechuan	533	A	2	1	{5}	0.94	—
Quiche	Mayan	337	A	2	1	{6}	0.97	1.00
			P	2	1	{5}	0.94	1.00
Reyesano	Pano-Tacanan	2997	A	3	2	{4, 1}	0.86	0.00
			P	3	1	{4}	0.96	0.00
Russian	Indo-European	340	A	2	1	{9}	1.00	—
<i>Shughni</i>	Indo-European	2885	A	1	1	{6}	—	—
<i>Sirionó</i>	Tupian	2476	A	1	1	{9}	—	—
			P	1	1	{7}	—	—
<i>Slovene</i>	Indo-European	2447	A	1	1	{7}	—	—
Swahili	Benue-Congo	361	A	3	1	{5}	0.99	1.00
			P	3	1	{5}	0.99	1.00
Tamashek (Burkina Faso)	Berber	3042	A	2	2	{4, 3}	0.00	—
Tamashek (Mali)	Berber	3040	A	2	2	{4, 3}	0.00	—
Tamazight (Ayt Ndhir)	Berber	571	A	2	2	{4, 3}	0.00	—
<i>Tapirapé</i>	Tupian	3261	A	1	1	{7}	—	—
			P	1	1	{5}	—	—
<i>Tenetebara</i>	Tupian	3269	A	1	1	{6}	—	—
Teprehuan (Southeastern)	Uto-Aztecan	2544	A	3	2	{3, 1}	0.67	0.93
			P	3	1	{5}	0.99	0.93
Teso	Nilotic	2548	A	3	2	{5, 1}	0.95	0.00
			P	3	2	{3, 1}	0.67	0.00
Thulung (Mukli)	Sino-Tibetan	667	A	6	5	{4, 2, 1, 1, 1}	0.59	0.00
			P	6	5	{4, 3, 1, 1, 1}	0.70	0.00
Tirmaga	Surmic	1414	A	5	3	{7, 3, 1}	1.00	0.92
			P	5	1	{4}	0.99	0.92
Tobati	Austronesian	2628	A	2	1	{4}	0.88	1.00
			P	2	1	{6}	0.97	1.00
Turkana	Nilotic	2641	A	3	2	{4, 1}	0.86	0.00
			P	3	2	{3, 1}	0.67	0.00
Turkish	Turkic	502	A	3	1	{5}	0.99	—
Tuva	Turkic	387	A	3	1	{6}	1.00	—
Tzutujil	Mayan	388	A	2	1	{6}	0.97	1.00
			P	2	1	{5}	0.94	1.00
Udihe	Tungusic	2657	A	2	2	{7, 6}	0.00	—
Udmurt	Uralic	679	A	2	2	{6, 5}	0.00	—
<i>Usan</i>	Madang	393	P	1	1	{6}	—	—
Wambule	Sino-Tibetan	2865	A	3	3	{10, 2, 2}	0.98	0.00
			P	3	3	{9, 2, 1}	0.97	0.00
<i>Xingú Asurini</i>	Tupian	3250	A	1	1	{8}	—	—
			P	1	1	{2}	—	—

(Appendix table continues)

LANGUAGE	STOCK	LANG ID	CAT	POS	POS	PARTITION	ALIGN	FEAT
				AVL	USED		INDEX	COH
Yagaria	Eastern Highlands	2869	A	3	1	{7}	1.00	1.00
			P	3	1	{8}	1.00	1.00
Yakkha	Sino-Tibetan	2996	A	13	7	{1, 1, 1, 1, 1, 1, 1}	0.00	0.00
			P	13	7	{1, 1, 1, 1, 1, 1, 1}	0.00	0.00
Yamphu	Sino-Tibetan	637	A	9	6	{3, 2, 2, 1, 1, 1}	0.47	0.00
			P	9	7	{3, 2, 2, 1, 1, 1, 1}	0.25	0.00
Yucatec	Mayan	682	A	5	2	{4, 2}	0.97	0.94
			P	5	2	{6, 6}	1.00	0.94
Zuni <sup>26</sup>	Zuni	429	A	2	1	{1}	—	—
			P	2	1	{1}	—	—

## REFERENCES

- ACKERMAN, FARRELL, and ROBERT MALOUF. 2013. Morphological organization: The low conditional entropy conjecture. *Language* 89(3).429–64. DOI: 10.1353/lan.2013.0054.
- ANDERSON, STEPHEN R. 1980. On the development of morphology from syntax. *Historical morphology*, ed. by Jacek Fisiak, 51–70. The Hague: Mouton. DOI: 10.1515/9783110823127.
- ANDERSON, STEPHEN R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511586262.
- ARNOTT, D. W. 1970. *The nominal and verbal systems of Fula*. Oxford: Oxford University Press.
- ARREGI, KARLOS, and ANDREW NEVINS. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout*. Berlin: Springer.
- BAAYEN, R. HARALD. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- BAKER, MARK. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16(3).373–415. Online: <https://www.jstor.org/stable/4178442>.
- BERGSMAN, WICHER. 2013. A bias-correction for Cramér's *V* and Tschuprow's *T*. *Journal of the Korean Statistical Society* 42(3).323–28. DOI: 10.1016/j.jkss.2012.10.002.
- BICKEL, BALTHASAR. 2011. Grammatical relations typology. *The Oxford handbook of linguistic typology*, ed. by Jae Jung Song, 399–444. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199281251.013.0020.
- BICKEL, BALTHASAR. 2015. Distributional typology: Statistical inquiries into the dynamics of linguistic diversity. *The Oxford handbook of linguistic analysis*, 2nd edn., ed. by Bernd Heine and Heiko Narrog, 901–23. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199677078.013.0046.
- BICKEL, BALTHASAR; GOMA BANJADE; MARTIN GAENZSLE; ELENA LIEVEN; NETRA PRASAD PAUDYAL; ICHCHHA PURNA RAI; MANOJ RAI; NOVEL KISHORE RAI; and SABINE STOLL. 2007. Free prefix ordering in Chintang. *Language* 83(1).43–73. DOI: 10.1353/lan.2007.0002.
- BICKEL, BALTHASAR, and JOHANNA NICHOLS. 2007. Inflectional morphology. *Language typology and syntactic description, vol. 3: Grammatical categories and the lexicon*, 2nd edn., ed. by Timothy Shopen, 169–240. Cambridge: Cambridge University Press.
- BICKEL, BALTHASAR; JOHANNA NICHOLS; TARAS ZAKHARKO; ALENA WITZLACK-MAKAREVICH; KRISTINE A. HILDEBRANDT; MICHAEL RIESSLER; LENNARD BIERKANDT; FERNANDO ZÚÑIGA; and JOHN B. LOWE. 2017. *The AUTOTYP typological databases*. Online: <https://github.com/autotyp/autotyp-data/tree/0.1.0>.
- BICKEL, BALTHASAR; SABINE STOLL; MARTIN GAENZSLE; N. K. RAI; ELENA LIEVEN; GOMA BANJADE; TOYA N. BHATTA; et al. 2017. Audiovisual corpus of the Chintang language. Online: <http://dobes.mpi.nl/>.
- BICKEL, BALTHASAR; ALENA WITZLACK-MAKAREVICH; KAMAL K. CHOUDHARY; MATTHIAS SCHLESEWSKY; and INA BORNKESSEL-SCHLESEWSKY. 2015. The neurophysiol-

<sup>26</sup> Zuni is excluded from the paradigmatic alignment bias calculation because just one agreement marker was extracted for each of the A and P roles; other languages are excluded because only one available position was extracted from the database.

- ogy of language processing shapes the evolution of grammar: Evidence from case marking. *PLOS ONE* 10(8):e0132819. DOI: 10.1371/journal.pone.0132819.
- BICKEL, BALTHASAR; TARAS ZAKHARKO; LENNART BIERKANDT; and ALENA WITZLACK-MAKAREVICH. 2014. Semantic role clustering: An empirical assessment of semantic role types in non-default case assignment. *Studies in Language* 38(3):485–511. DOI: 10.1075/sl.38.3.03bic.
- BICKEL, BALTHASAR, and FERNANDO ZÚÑIGA. 2017. The ‘word’ in polysynthetic languages: Phonological and syntactic challenges. *The Oxford handbook of polysynthesis*, ed. by Michael Fortescue, Marianne Mithun, and Nicholas Evans, 158–85. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199683208.013.52.
- BRESNAN, JOAN; SHIPRA DINGARE; and CHRISTOPHER D. MANNING. 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. *Proceedings of the LFG01 Conference*, 1–20. Online: <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/6/pdfs/lfg01bresnanetal.pdf>.
- BÜRKNER, PAUL-CHRISTIAN. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80:1–28. DOI: 10.18637/jss.v080.i01.
- BYBEE, JOAN L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- CABALLERO, GABRIELA. 2010. Scope, phonology and morphology in an agglutinating language: Choguita Rarámuri (Tarahumara) variable suffix ordering. *Morphology* 20(1): 165–204. DOI: 10.1007/s11525-010-9147-4.
- CHEMLA, EMMANUEL; TOBEN H. MINTZ; SAVITA BERNAL; and ANNE CHRISTOPHE. 2009. Categorizing words using ‘frequent frames’: What cross-linguistic analyses reveal about distributional acquisition strategies. *Developmental Science* 12(3):396–406. DOI: 10.1111/j.1467-7687.2009.00825.x.
- CHRISTIANSEN, MORTEN H., and NICK CHATER. 2008. Language as shaped by the brain. *Behavioral and Brain Sciences* 31(5):489–509. DOI: 10.1017/S0140525X08004998.
- CRIBARI-NETO, FRANCISCO, and ACHIM ZEILEIS. 2010. Beta regression in R. *Journal of Statistical Software* 34(1):1–24. DOI: 10.18637/jss.v034.i02.
- CRYSMANN, BERTHOLD. 2017. Inferential-realizational morphology without rule blocks: An information-based approach. *Defaults in morphological theory*, ed. by Nikolas Gisborne and Andrew Hippisley, 182–213. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198712329.003.0008.
- CRYSMANN, BERTHOLD, and OLIVIER BONAMI. 2016. Variable morphotactics in information-based morphology. *Journal of Linguistics* 52(2):311–74. DOI: 10.1017/S0022226715000018.
- CULBERTSON, JENNIFER; PAUL SMOLENSKY; and GÉRALDINE LEGENDRE. 2012. Learning biases predict a word order universal. *Cognition* 122(3):306–29. DOI: 10.1016/j.cognition.2011.10.017.
- DEDIU, DAN; RICK JANSSEN; and SCOTT R. MOISIK. 2017. Language is not isolated from its wider environment: Vocal tract influences on the evolution of speech and language. *Language & Communication* (Special issue: *The multimodal origins of linguistic communication*, ed. by Sławomir Waciewicz and Przemysław Żywiecżyński) 54:9–20. DOI: 10.1016/j.langcom.2016.10.002.
- DELANCEY, SCOTT. 1981. An interpretation of split ergativity and related patterns. *Language* 57(3):626–57. DOI: 10.2307/414343.
- DIESING, MOLLY; DUŠICA FILIPOVIĆ ĐURĐEVIĆ; and DRAGA ZEC. 2009. Clitic placement in Serbian: Corpus and experimental evidence. *The fruits of empirical linguistics, vol. 2: Product*, ed. by Susanne Winkler and Sam Featherstone, 59–74. Berlin: Mouton de Gruyter.
- DOORNENBAL, MARIUS. 2009. *A grammar of Bantawa*. Meteren: Netherlands Graduate School of Linguistics.
- DRYER, MATTHEW S. 2013. Order of subject, object and verb. *The world atlas of language structures online*, ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: <http://wals.info/chapter/81>.
- DUÑABEITIA, JON ANDONI; ITZIAR LAKA; MANUEL PEREA; and MANUEL CARREIRAS. 2009. Is *Milkman* a superhero like *Batman*? Constituent morphological priming in compound words. *European Journal of Cognitive Psychology* 21(4):615–40. DOI: 10.1080/09541440802079835.

- EBERT, KAREN H. 1987. Grammatical marking of speech act participants in Tibeto-Burman. *Journal of Pragmatics* 11(4).473–82. DOI: 10.1016/0378-2166(87)90090-7.
- EMBICK, DAVID. 2015. *The morpheme: A theoretical introduction*. Berlin: De Gruyter Mouton.
- EMBICK, DAVID, and ROLF NOYER. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4).555–95. DOI: 10.1162/002438901753373005.
- EMBICK, DAVID, and ROLF NOYER. 2007. Distributed morphology and the syntax–morphology interface. *The Oxford handbook of linguistic interfaces*, ed. by Gillian Ramchand and Charles Reiss, 289–324. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199247455.013.0010.
- FERMINO, JESSIE LITTLE DOE. 2000. *An introduction to Wampanoag grammar*. Cambridge, MA: MIT dissertation. Online: <http://hdl.handle.net/1721.1/8740>.
- FERRARI, SILVIA, and FRANCISCO CRIBARI-NETO. 2004. Beta regression for modelling rates and proportions. *Journal of Applied Statistics* 31(7).799–815. DOI: 10.1080/0266476042000214501.
- FOLEY, WILLIAM A., and ROBERT D. VAN VALIN, JR. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- GAGNÉ, CHRISTINA L.; THOMAS L. SPALDING; LAUREN FIGUEREDO; and ALLISON C. MULLALY. 2009. Does snow man prime plastic snow?: The effect of constituent position in using relational information during the interpretation of modifier-noun phrases. *The Mental Lexicon* 4(1).41–76. DOI: 10.1075/ml.4.1.03gag.
- GIVÓN, TALMY. 1971. Historical syntax and synchronic morphology: An archaeologist's field trip. *Chicago Linguistic Society* 7(1).394–415.
- GODDARD, IVES. 2000. The historical origins of Cheyenne inflections. *Papers of the 31st Algonquian Conference*, ed. by John Nichols, 77–129. Winnipeg: University of Manitoba.
- GODDARD, IVES, and KATHLEEN J. BRAGDON. 1988. *Native writings in Massachusetts*. Philadelphia: American Philosophical Society.
- GOOD, JEFF. 2016. *The linguistic typology of templates*. Cambridge: Cambridge University Press.
- GOOD, JEFF, and ALAN C. L. YU. 2005. Morphosyntax of two Turkish subject pronominal paradigms. *Clitic and affix combinations: Theoretical perspectives*, ed. by Lorie Heggie and Francisco Ordóñez, 315–41. Amsterdam: John Benjamins.
- GORMAN, KYLE, and DANIEL EZRA JOHNSON. 2013. Quantitative analysis. *The Oxford handbook of sociolinguistics*, ed. by Robert Bayley, Richard Cameron, and Ceil Lucas, 214–40. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199744084.013.0011.
- GREEN, CHRISTOPHER R., and MICHELLE E. MORRISON. 2016. Somali wordhood and its relationship to prosodic structure. *Morphology* 26(1).3–32. DOI: 10.1007/s11525-015-9268-x.
- GUILLAUME, ANTOINE. 2009. Hierarchical agreement and split intransitivity in Reyesano. *International Journal of American Linguistics* 75(1).29–48. DOI: 10.1086/598202.
- HALLE, MORRIS, and ALEC MARANTZ. 1993. Distributed morphology and the pieces of inflection. *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 111–76. Cambridge, MA: MIT Press.
- HARDY, G. H., and E. M. WRIGHT. 2008. *An introduction to the theory of numbers*. 6th edn. Oxford: Oxford University Press.
- HARRIS, ALICE C. 2002. *Endoclititics and the origins of Udi morphosyntax*. Oxford: Oxford University Press.
- HARRIS, ALICE C. 2017. *Multiple exponence*. Oxford: Oxford University Press.
- HAWKINS, JOHN A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511554285.
- HAWKINS, JOHN A. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press.
- HAY, JENNIFER. 2002. From speech perception to morphology: Affix ordering revisited. *Language* 78(3).527–55. DOI: 10.1353/lan.2002.0159.
- HIMMELMANN, NIKOLAUS P. 2014. Asymmetries in the prosodic phrasing of function words: Another look at the suffixing preference. *Language* 90(4).927–60. DOI: 10.1353/lan.2014.0105.

- HOCKETT, CHARLES F. 1966. What Algonquian is really like. *International Journal of American Linguistics* 32(1).59–73. DOI: 10.1086/464880.
- HYMAN, LARRY M. 2003. Suffix ordering in Bantu: A morphocentric approach. *Yearbook of Morphology* 2002.245–81. DOI: 10.1007/0-306-48223-1\_8.
- INKELAS, SHARON. 1993. Nimboran position class morphology. *Natural Language and Linguistic Theory* 11.559–624. DOI: 10.1007/BF00993014.
- JACQUES, GUILLAUME; AIMÉE LAHAUSSOIS; BOYD MICHAÏLOVSKY; and DHAN BAHADUR RAI. 2012. An overview of Khaling verbal morphology. *Language and Linguistics* 13(6).1095–1170.
- JULIEN, MARIT. 2002. *Syntactic heads and word formation*. Oxford: Oxford University Press.
- KEMMERER, DAVID. 2012. The cross-linguistic prevalence of SOV and SVO word orders reflects the sequential and hierarchical representation of action in Broca's area. *Language and Linguistics Compass* 6(1).50–66. DOI: 10.1002/inc3.322.
- KIM, YUNI. 2010. Phonological and morphological conditions on affix order in Huave. *Morphology* 20(1).133–63. DOI: 10.1007/s11525-010-9149-2.
- LEVSHINA, NATALIA. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.
- LUUTONEN, JORMA. 1997. *The variation of morpheme order in Mari declension*. Helsinki: Suomalais-Ugrilainen Seura.
- MACDONALD, MARYELLEN C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4:226. DOI: 10.3389/fpsyg.2013.00226.
- MANOVA, STELA, and MARK ARONOFF. 2010. Modeling affix order. *Morphology* 20(1).109–31. DOI: 10.1007/s11525-010-9153-6.
- MANSFIELD, JOHN BASIL. 2015. Morphotactic variation, prosodic domains and the changing structure of the Murrinhpatha verb. *Asia-Pacific Language Variation* 1(2).163–89. DOI: 10.1075/aplv.1.2.03man.
- MANSFIELD, JOHN BASIL. 2019. *Murrinhpatha morphology and phonology*. Berlin: De Gruyter Mouton.
- MINTZ, TOBEN H. 2002. Category induction from distributional cues in an artificial language. *Memory & Cognition* 30(5).678–86. DOI: 10.3758/BF03196424.
- MINTZ, TOBEN H. 2003. Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 90(1).91–117. DOI: 10.1016/S0010-0277(03)00140-9.
- MINTZ, TOBEN H.; FELIX HAO WANG; and JIA LI. 2014. Word categorization from distributional information: Frames confer more than the sum of their (bigram) parts. *Cognitive Psychology* 75.1–27. DOI: 10.1016/j.cogpsych.2014.07.003.
- MORAN, STEVEN; DAMIÁN E. BLASI; ROBERT SCHIKOWSKI; AYLIN C. KÜNTAY; BARBARA PFEILER; SHANLEY ALLEN; and SABINE STOLL. 2018. A universal cue for grammatical categories in the input to children: Frequent frames. *Cognition* 175.131–40. DOI: 10.1016/j.cognition.2018.02.005.
- NAPOLI, DONNA JO; NATHAN SANDERS; and REBECCA WRIGHT. 2014. On the linguistic effects of articulatory ease, with a focus on sign languages. *Language* 90(2).424–56. DOI: 10.1353/lan.2014.0026.
- NERCESIAN, VERÓNICA. 2014. Wordhood and the interplay of linguistic levels in synthetic languages: An empirical study on Wichi (Mataguayan, Gran Chaco). *Morphology* 24(3).177–98. DOI: 10.1007/s11525-014-9239-7.
- NICHOLS, JOHANNA. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- NORDLINGER, RACHEL. 2015. Inflection in Murrinh-Patha. *The Oxford handbook of inflection*, ed. by Matthew Baerman, 491–519. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199591428.013.21.
- NOYER, ROLF. 1997. *Features, positions and affixes in autonomous morphological structure*. New York: Garland.
- PASTER, MARY. 2009. Explaining phonological conditions on affixation: Evidence from suppletive allomorphy and affix ordering. *Word Structure* 2(1).18–37. DOI: 10.3366/E1750124509000282.
- PLAG, INGO, and HARALD BAAYEN. 2009. Suffix ordering and morphological processing. *Language* 85(1).109–52. DOI: 10.1353/lan.0.0087.
- RICE, KEREN. 2000. *Morpheme order and semantic scope: Word formation in the Athapaskan verb*. Cambridge: Cambridge University Press.

- RICE, KEREN. 2011. Principles of affix ordering: An overview. *Word Structure* 4(2).169–200.
- ROBERTS, JOHN R. 1987. *Amele*. London: Croom Helm.
- RYAN, KEVIN M. 2010. Variable affix order: Grammar and learning. *Language* 86(4).758–91. DOI: 10.1353/lan.2010.0032.
- SAGOT, BENOÎT, and GÉRALDINE WALTHER. 2011. Non-canonical inflection: Data, formalisation and complexity measures. *Systems and frameworks for computational morphology* (Communications in computer and information science), ed. by Cerstin Mahlow and Michael Piotrowski, 23–45. Berlin: Springer. DOI: 10.1007/978-3-642-23138-4\_3.
- SCHACHTER, PAUL, and FE T. OTANES. 1983. *Tagalog reference grammar*. Berkeley: University of California Press.
- SCHAKOW, DIANA. 2015. *A grammar of Yakkha*. Berlin: Language Science. DOI: 10.17169/langsci.b66.106.
- SCHIKOWSKI, ROBERT. 2014. Chintang sketch grammar. Zurich: University of Zurich, ms.
- SCHIKOWSKI, ROBERT; NETRA PRASAD PAUDYAL; and BALTHASAR BICKEL. 2015. Flexible valency in Chintang. *Valency classes in the world's languages*, ed. by Bernard Comrie and Andrej Malchukov, 669–707. Berlin: De Gruyter Mouton.
- SCHWENTER, SCOTT A., and RENA TORRES CACOULOS. 2014. Competing constraints on the variable placement of direct object clitics in Mexico City Spanish. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics* 27(2).514–36. DOI: 10.1075/resla.27.2.13sch.
- SEIFART, FRANK; JAN STRUNK; SWINTHA DANIELSEN; IREN HARTMANN; BRIGITTE PAKENDORF; SØREN WICHMANN; ALENA WITZLACK-MAKAREVICH; NIVJA H. DE JONG; and BALTHASAR BICKEL. 2018. Nouns slow down speech across structurally and culturally diverse languages. *Proceedings of the National Academy of Sciences* 115(22).5720–25. DOI: 10.1073/pnas.1800708115.
- SHANNON, CLAUDE E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27(3).379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.
- SIEWIERSKA, ANNA. 2004. *Person*. Cambridge: Cambridge University Press.
- SIMPSON, JANE. 2007. Expressing pragmatic constraints on word order in Warlpiri. *Architectures, rules, and preferences: Variations on themes by Joan W. Bresnan*, ed. by Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling, and Chris Manning, 403–27. Stanford, CA: CSLI Publications.
- SIMPSON, JANE, and ILANA MUSHIN. 2008. Clause-initial position in four Australian languages. *Discourse and grammar in Australian languages*, ed. by Ilana Mushin and Brett Baker, 25–57. Amsterdam: John Benjamins.
- SIMPSON, JANE, and M. WITHGOTT. 1986. Pronominal clitic clusters and templates. *The syntax of pronominal clitics*, ed. by Hagit Borer, 147–74. New York: Academic Press. DOI: 10.1163/9789004373150\_008.
- SMOLENSKY, PAUL, and GÉRALDINE LEGENDRE. 2006. *The harmonic mind: From neural computation to optimality-theoretic grammar*. Cambridge, MA: MIT Press.
- SPEELMAN, DIRK. 2014. Logistic regression: A confirmatory technique for comparisons in corpus linguistics. *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, ed. by Dylan Glynn and Justyna A. Robinson, 487–533. Amsterdam: John Benjamins.
- SPENCER, ANDREW. 2003. Putting some order into morphology: Reflections on Rice (2000) and Stump (2001). *Journal of Linguistics* 39(3).621–46. DOI: 10.1017/S002222670302123.
- STOLL, SABINE; BALTHASAR BICKEL; and JEKATERINA MAZARA. 2017. The acquisition of polysynthetic verb forms in Chintang. *The Oxford handbook of polysynthesis*, ed. by Michael Fortescue, Marianne Mithun, and Nicholas Evans, 495–516. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199683208.013.28.
- STUMP, GREGORY T. 1997. Template morphology and inflectional morphology. *Yearbook of Morphology* 1996.217–41. DOI: 10.1007/978-94-017-3718-0\_12.
- STUMP, GREGORY T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.
- SZMRECSANYI, BENEDIKT. 2006. *Morphosyntactic persistence in spoken English: A corpus study at the intersection of variationist sociolinguistics*. Berlin: Mouton de Gruyter.



- TRAVIS, CATHERINE E. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19(2).101–35. DOI: 10.1017/S0954394507070081.
- TROMMER, JOCHEN. 2003. The interaction of morphology and syntax in affix order. *Yearbook of Morphology* 2002.283–324. DOI: 10.1007/0-306-48223-1\_9.
- VAN EGMOND, MARIE-ELAINE. 2012. *Enindhilyakwa phonology, morphosyntax and genetic position*. Sydney: University of Sydney dissertation.
- VON HUMBOLDT, WILHELM. 1836. *Über die Verschiedenheit des menschlichen Sprachbaus und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechtes*. Berlin: Dümmler.
- WIDMER, MANUEL; SANDRA AUDERSET; JOHANNA NICHOLS; PAUL WIDMER; and BALTHASAR BICKEL. 2017. NP recursion over time: Evidence from Indo-European. *Language* 93(4).799–826. DOI: 10.1353/lan.2017.0058.
- WITZLACK-MAKAREVICH, ALENA; TARAS ZAKHARKO; LENNART BIERKANDT; FERNANDO ZÚÑIGA; and BALTHASAR BICKEL. 2016. Decomposing hierarchical alignment: Co-arguments as conditions on alignment and the limits of referential hierarchies as explanations in verb agreement. *Linguistics* 54(3).531–61. DOI: 10.1515/ling-2016-0011.

Mansfield  
 School of Languages and Linguistics  
 University of Melbourne  
 Parkville VIC 3010, Australia  
 [john.mansfield@unimelb.edu.au]  
 [sabine.stoll@uzh.ch] (Stoll)  
 [balthasar.bickel@uzh.ch] (Bickel)

[Received 15 March 2018;  
 revision invited 6 September 2018;  
 revision received 29 January 2019;  
 revision invited 2 April 2019;  
 revision received 26 May 2019;  
 revision invited 4 September 2019;  
 revision received 27 October 2019;  
 accepted 16 December 2019]