Chapter 3

Deconstructing exuberant exponence in Batsbi

Berthold Crysmann

Université de Paris, Laboratoire de linguistique formelle, CNRS

In this article, I shall discuss "exuberant exponence" in Batsbi (Harris 2009), an extreme case of extended exponence where identical gender-number markers can surface multiple times within the same word, subject to the presence of certain triggering stems or affixes. I shall also evaluate in some detail the challenge the Batsbi data pose for extant formal theories of inflection and show that these challenges cut across the divide between lexical and inferential theories. In the analysis, I shall highlight the dependent nature of the agreement exponents and propose a formal account that draws crucially on two central properties of Information-based Morphology, namely the recognition of many-to-many relations at the most fundamental level of description, and the possibility to extract (partial) generalisations over rules by means of cross-classifying inheritance hierarchies. As a result, cross-classification of agreement rule types with those for the triggering stems and affixes will capture the dependent nature directly, while at the same time ensuring the reuse of inflectional resources. Thus, the decomposition of Batsbi exuberant exponence improves considerably over a pure word-based approach and emphasises the need to afford both atomistic and holistic views within a theory of inflection.

1 Introduction

Ever since Matthews (1972), extended (or multiple) exponence has been one of the core phenomena highlighting the one-to-many nature of inflectional morphology (see Harris 2017 for a typological survey). In this chapter, I shall discuss exuberant exponence in Batsbi (Harris 2009), an extreme case of extended exponence, where one and the same morphosyntactic property may end up being

Berthold Crysmann. 2021. Deconstructing exuberant exponence in Batsbi. In Berthold Crysmann & Manfred Sailer (eds.), *One-to-many relations in morphology, syntax, and semantics*, 53–84. Berlin: Language Science Press. DOI: 10.5281/zenodo.4729793

marked over and over again within a word. Outside Batsbi, the phenomenon has been reported for a variety of languages, including Archi, Khinalug, Chamalal (see Harris 2009 for an extended list).

Exuberant exponence in Batsbi is manifest in gender/number agreement on verbs, giving rise to up to four realisations of agreement with the same argument, the absolutive. What is more, the shape of the exponents across multiple realisations stays the same.

(1) **y**-ox-**y**-∅-o-**y**-anŏ
CM-rip-CM-TR-PRES-CM-EVID1
'Evidently she ripped it.' (Harris 2009: 277)

What makes exuberant exponence particularly interesting from the viewpoint of formal grammar is that the phenomenon can serve as a stress-test for current theories of inflectional morphology. First, exuberant exponence will be less troublesome for theories that fully embrace extended exponence as a basic property of inflectional morphology, rather than providing limited workarounds on the basis of an essentially morphemic model. Second, the identity of exponents observed in Batsbi calls for inflectional models that provide a notion of resource reusability. Third, as I shall discuss below, the presence of agreement markers is dependent on adjacent triggering stems and suffixes, which suggests that agreement markers cannot be derived on their own, but rather compose with the affixes that license their occurrence into inflectional constructions. I shall argue more specifically that the dependent nature of Batsbi exuberant exponence calls for a model of morphology that addresses the many-to-many nature of inflection at the most basic level, a property characteristic of the framework adopted here.

The presentation of the empirical facts about Batsbi exuberant exponence is based on the original paper by Harris (2009). Thus, this paper aims at making contributions in two areas: first explore in more detail the implications of the data for different incarnations of inferential-realisational and inferential-lexical approaches, and second, provide a fully formalised treatment of this challenging case of dependent multiple exponence within the framework of Information-based Morphology (=IbM; Crysmann & Bonami 2016).

The chapter is organised as follows: in Section 2, I shall rehearse the basic empirical data concerning Batsbi exuberant exponence, starting with the inventory of (productive) gender markers, followed by a discussion of class marking on stems as well as affixal material. Section 3 will serve to evaluate extant theories of inflection with respect to their capability to address the phenomenon at hand, taking as a starting point the typology developed in Stump (2001): While incremental theories prove to be inadequate, a somewhat striking observation is that exuberant exponence does not distinguish between lexical-realisational and incremental-realisational models as a class, but is rather sensitive to details of formal expressivity of the concrete theory.

Section 5 will finally provide an analysis within the framework of Information-based Morphology (henceforth: IbM), an inferential-realisational model of morphology cast entirely in terms of inheritance hierarchies of typed feature structures. I shall provide a brief sketch of IbM and then show how cross-classification in monotonic inheritance hierarchies is well-suited to capture reuse of form and the dependent nature of exuberant exponence at the same time.

2 Data

2.1 Properties of class marking in Batsbi

Batsbi has a rather elaborate gender system, distinguishing eight gender categories, each with singular and plural forms, out of which at least five are productive, while the following three are not, according to Corbett (1991) and Holisky & Gagua (1994): genders IV (2 nouns), VIII (4), and VII (15). Lexical counts are indicated in parentheses.

Exponence of gender/class agreement is detailed in Table 3.1. As can be seen, /d/ is quite prevalent as an exponent, which is why Harris occasionally uses it as a representative for the entire set of class markers.

Table 3.1: Gender agreement markers in Batsbi

	SINGULAR	PLURAL
I	v	b
II	y	d
III	y	y
(IV)	b	b
V	d	d
VI	b	d
(VII)	b	y
(VIII)	d	y

Gender/number agreement is controlled by the absolutive argument, i.e. the S argument of intransitives, as witnessed in (2), and the O argument of transitives, as shown by (3).

(2) xen-go-ħ potl-i **d**-ek'-ĩ
tree-ALL-LOC leaf(d/d)-PL.ABS CM-fall-AOR
'The leaves of the tree were falling.' (Harris 2009: 274)

(3) pst'uyn-čo-v bader **d**-iy-ẽ
married.woman(y/y)-OBL-ERG child(d/d).ABS CM-do-AOR
'The (married) woman bore a child.' (Harris 2009: 274)

2.2 Class marking on stems

As we have seen in example (1) above, Batsbi class marking can surface multiple times within a word, and when it does, we always find the same exponents. However, as pointed out by Harris (2009), presence of class markers in this language is contingent on the right-adjacent marker: just as we may find words with multiple class markers, as in (1), we may equally find words showing a single marker, as in (2), (3) or (4), or even no overt class making at all, as e.g. in (5).

(4) oqus mot: k'edl-e-guy tat:-**b**-iy-ẽ
3SG.ERG bed(b/d).ABS wall-OBL-towards push-CM-TR-AOR
'S/he pushed the bed towards the wall.' (Harris 2009: 275)

(5) qan simind lapsdan matx ot'-ŏ
tomorrow corn(d/d).ABS to.dry sun(b/d) spread-FUT
'Tomorrow [they] will spread the corn in the sun to dry.' (idem)

Stems are one of the elements that may require or disallow left adjacent class markers: according to Harris (2009: fn. 23), 468 (21.53%) out of 2178 verbs in the dictionary by Kadagiʒe & Kadagiʒe (1984) feature a pre-radical class marker. While none of the stems in (4) or (5) appears to take a class marker to its immediate left, the verbs *ek'* 'fall' and *iy* 'do' in fact do, as illustrated in (2) and (3) above.

Holisky & Gagua (1994) note that some verbs distinguish the perfective from the imperfective stem by means of an agreement marker, contrasting, e.g. ***d****-ek'-ar* 'fall.PFV' with *ak'-ar* 'fall.IPFV'. Harris (2009) provides a list of minimal pairs, where lexical meaning is solely distinguished by presence of a pre-radical marker, including e.g. *ot:-ar* 'stand, stay' vs. ***d****-ot:-ar* 'pour into'. Thus, it appears that the presence vs. absence of a pre-radical agreement marker is lexically determined, i.e. it is a property of individual stems, or else of the entire lexeme. Choice of the shape of the marker, by contrast, is clearly an inflectional property.

2.3 Class marking on suffixes

Similar to pre-stem class markers, post-stem gender/number markers appear left-adjacent to certain triggering suffixes. These comprise the transitivity markers *-al* (INTR) and *-iy* (TR), as well as the affirmative and negative evidentiality markers *-anŏ* (EVID1.AFF) and *-a* (EVID1.NEG). Pre-stem and pre-suffixal class markers are controlled by the same argument, the absolutive, and their shape is identical. However, their presence is conditioned independently.

Table 3.2: Patterns of dependent class marking in Batsbi (Harris 2009)

STEM	(IN)TRANS	EVID
stem	∅	∅/-inŏ
d-stem	∅	∅/-inŏ
stem	∅	**d**-anŏ
d-stem	∅	**d**-anŏ
stem	**d**-iy/al	∅/-inŏ
d-stem	**d**-iy/al	∅/-inŏ
stem	**d**-iy/al	**d**-anŏ
d-stem	**d**-iy/al	**d**-anŏ

2.3.1 Intransitive marker *-al*

The basic function of the intransitive marker *-al* is to derive intransitives from transitives, as illustrated in (6).[1]

(6) a. p'erang-mak-aħŏ xalat y-opx-ŏ
shirt-on-LOC house.coat(y/y).ABS CM-put.on-PRS
'[She] puts on a house coat over her shirt.' (Harris 2009: 275)

b. sẽ yoħ taguš y-opx-y-al-in=ĕ
me.GEN girl(y/d).ABS beautifully CM-put.on-CM-INTR-AOR=&
'My daughter dressed beautifully and ...' (Harris 2009: 275)

When this marker is present, it is obligatorily accompanied by the class marker to its left. Presence of the post-radical marker is triggered independently of the stem, as shown by the contrast between (6a,b) and (7).

[1]Note that the intransitive marker can also be found with some intransitive bases, e.g. *ak'-**d**-al-ar* 'light up, catch fire' vs. *ak'-ar* 'burn, be alight'.

(7) psare(ħ) oc'-**v**-al-in-es ...
yesterday weigh.PFV-CM-INTR-AOR-1SG.ERG
'I (masculine) weighed yesterday ...' (Harris 2009: 275)

2.3.2 Transitive marker *-iy*

While the intransitive marker *-al* derives intransitives from transitive bases, the transitive marker *-iy* signals the opposite, namely transitives derived from intransitive bases.[2] Again, this marker is immediately preceded by the class marker, as illustrated in the examples in (8).

(8) a. don-e-v taylz-i **d**-ek'-**d**-iy-ẽ
horse(b/d)-OBL-ERG saddlebags(/d)[3]-PL.ABS CM-fall-CM-TR-AOR
'The horse threw off the saddlebags.' (Harris 2009: 274)

b. kuyrc'l-e-x qečqečnayrẽ daq'r-i lal-**d**-iy-ẽ
wedding-OBL-CON various food(d/d)-PL.ABS go-CM-TR-AOR
makaħŏ
above
'At the wedding [they] passed around various foods.' (idem)

2.3.3 Present evidential

The third suffixal marker that takes the class marker, again to its immediate left, is the present evidential marker *-anŏ*. According to Harris (2009), this marker productively combines with any lexeme. Compare the examples in (9): adding the present evidential to an example like (9a), with already two class markers (one triggered by the stem and one triggered by the transitive marker), adds a third instance of class marking, yielding a total of three exponents, as shown in (9b).

(9) a. k'ab **y**-ox-**y**-iy-ẽ
dress(y/y).ABS CM-rip-CM-TR-AOR
'[She] ripped the dress.' (Harris 2009: 277)

b. **y**-ox-**y**-∅-o-**y**-anŏ
CM-rip-CM-TR-PRS-CM-EVID1
'Evidently she ripped it.' (Harris 2009: 277)

[2] This marker may occasionally serve to distinguish transitives.

[3] 'Saddlebags' is a plurale tantum (Harris 2009: 274). Lacking an attested singular form, its gender could be any of II, V, or VI.

Again, class inflection of the evidential is independent of that of the stem, i.e. it is triggered by the present evidential, regardless of whether the stem is already marked with the gender marker, as in (9b), or not, as in (10).

(10) tet'-**d**-anŏ
cut-CM-EVID1
'Evidently s/he was cutting it.' (Holisky & Gagua 1994: 181)

The present evidential *-anŏ* (EVID1) contrasts with, e.g. the aorist evidential *-inŏ*, which never takes a gender/number marker.

2.4 Wordhood

The implications of exuberant exponence for morphological theory depend of course on the crucial question whether the relevant domain is morphology, i.e. whether we are dealing with complex words, or syntax. Harris (2009) provides extensive tests showing that we are indeed confronted with massive extended exponence within a single word, rather than agreement across several syntactically independent words. This is even more important given that most markers involved here used to be independent words diachronically, e.g. the evidential marker derives from the verb 'to be'.

Regarding the status of class markers, Harris (2009) provides five tests in total.[4] I shall give a brief description of the tests, and summarise the results, which uniformly point towards the affixal status of the class markers (see Table 3.3 for a summary of the results, and Harris 2009: sec. 5 for details).

Agreement controller: Establishes whether auxiliary or evidential markers share an argument structure with the verb: true auxiliaries behave like intransitives (regardless of main verb), evidentials reflect the main verb's transitivity, suggesting bound status.

Intervention: Two related tests based on the possibility for intervention of negative marker and clitic conjunction: the possibility for intervention is independently established for auxiliaries, yet all markers under discussion uniformly prohibit intervention.

[4]Harris (2009) presents a total of seven tests, two of which are confined to the status of person/number markers. These markers may incidentally be controlled by the same argument, which leads Harris to regard them as yet another instance of (partial) exuberant agreement. However, given that the controllers need not be the same (see Harris 2009: ex. 33)), I shall rather treat this as accidental, and thus ignore person/number marking for the purposes of this paper.

Coordination & Gapping: Two tests that assess whether or not markers can be suppressed in coordinate structures. While auxiliaries and main verbs can be elided in the second conjunct, transitive markers and evidentials cannot.

Table 3.3: Tests for word vs. affix status (Harris 2009)

Test	TR	INTR	EVID1	AUX
Agreement trigger			aff	wd
Intervention (neg)	aff	aff	aff	wd
Intervention (clitic)			aff	wd
Conjoining	aff	aff	aff	wd
Gapping	aff			wd

To summarise, the evidence Harris (2009) provides robustly points in the same direction, namely that transitivity markers and evidential markers are bound affixes. Therefore, the issue of exuberant exponence and the dependent nature of the class markers are to be addressed in the domain of morphology rather than relegating them to syntax.

3 Discussion

Exuberant exponence can probably be regarded as just another case of extended (or multiple) exponence, so we would expect theories that embrace the notion of many-to-many relations between function and form to outperform those which picture morphology in terms of (classical) morphemes. This is indeed the line of argumentation put forth by Harris (2009). In her article, she discusses the theoretical significance of extended exponence in general and exuberant exponence in particular and confronts the Batsbi facts with claims made by various theoretical frameworks. In particular, she observes that incremental theories are uniformly hard pressed to cover the empirical patterns, since these approaches assume that morphological operations must always add information, as in the lexical-incremental theory of Wunderlich & Fabri (1995), or must always express information not yet expressed, as in the inferential-incremental approach of Steele (1995).

3.1 Implications for lexical-realisational theories

Harris (2009) already discusses in some depth the implications of the Batsbi data for two instances of Distributed Morphology, a lexical-realisational theory in terms of the typology of morphological theories proposed by Stump (2001). She shows convincingly that the theory of primary and secondary exponence advanced by Noyer (1992) restricts extended exponence to maximally two occurrences, which makes it impossible to capture the Batsbi data, even though it may be adequate for Berber and Arabic, the languages Noyer based his theory on.

In a paper on extended exponence in German, Müller (2007) suggests to complement the theory of impoverishment (used in Distributed Morphology, Halle & Marantz 1993, to account for syncretism) with a theory of enrichment, in order to facilitate the treatment of extended exponence. In the interest of limiting the formal complexity of a system that recognises both deletion and insertion rules, he suggests that enrichment may only redundantly add features already present. As shown by Harris (2009), enrichment rules indeed make it possible for a lexical-realisational theory such as DM to cover the Batsbi data. The criticism she raises against the theory of enrichment is more of a conceptual nature, essentially stating that lexical theories are not well-equipped to capture the relevant generalisations directly, but rather force the surface patterns into a morphemic mould.

While I concur with Harris's general assessment of the two DM approaches, it is still worth noting that the problems faced by Noyer (1992) and by Müller (2007), are of an entirely different nature: while Müller's approach can indeed be criticised for favouring a morphemic ideal and deriving exuberant exponence by means of a "workaround", as argued convincingly in Harris (2009), it is equally clear that the theory of enrichment meets at least the criterion of weak generative capacity, unlike Noyer (1992). One might even suggest that the division between a morphemic core and enrichment could be motivated by considerations of what is or could be considered typologically canonical or unmarked. Noyer's theory contrasts sharply with that of Müller: his theory fails on grounds of weak capacity, i.e. it cannot even describe the set of acceptable surface words. What is more, the reason for this failure is located not at the level of the theory, where one might just drop some universal claim in favour of a language-particular constraint, but rather it is implemented at the level of the underlying logic of feature discharge, meaning there is just no chance of repair. To summarise, exuberant exponence falsifies Noyer's theory of feature discharge, while Müller's theory appears to be flexible enough to describe the facts.

The observation that there is no clear alignment with general properties of the approaches, but rather a strong dependence on the details of implementation suggests that a typology of morphological theories can only give a coarse indication of the analytical properties of a theory and therefore still needs to be complemented by careful investigation of the formal properties of the individual approaches.

3.2 Batsbi and inferential-realisational theories

In contrast to both morpheme-based (=lexical) and incremental theories, inferential-realisational theories generally embrace extended exponence as a recurrent property in inflectional systems. However, it seems that this very fact has led Harris (2009) to take for granted that every approach within this family of theories will be able to capture the empirical patterns. While there certainly is no general obstacle, we shall see in this section that not all word-and-paradigm theories are equally well-equipped to account for the Batsbi data in an insightful and maximally general fashion. To illustrate this point, I shall briefly discuss A-morphous Morphology (=AM; Anderson 1992) and Paradigm Function Morphology (=PFM; Stump 2001) and argue that it is important to submit to further scrutiny the architectural decisions and the formal devices offered by each theory.

A-morphous Morphology (AM) organises inflectional rules into a system of ordered rule blocks that is used to derive affix order. While there is preemption within rule blocks, by way of extrinsic rule ordering, preemption does not generally apply across different blocks, thereby making it possible in principle that a morphosyntactic property may get expressed more than once. However, AM does not provide any device permitting reuse of resources across different rule blocks. Thus, while extended exponence or even exuberant exponence per se is not a problem at all for Anderson's model, the absence of, e.g. rules of referral makes it difficult to capture the generalisation that exponents of gender marking are indeed identical across different surface positions in the word. Thus in addition to massive duplication of gender-marking rules across different rule blocks, surface identity is pictured as entirely accidental.

Paradigm Function Morphology also builds on a system of extrinsically ordered rule blocks and it equally limits rule competition to rules within the same block. In contrast to AM, however, PFM does provide rules of referral, either in terms of rules of referral to an ordered rule block (cf. Stump 1993), or by means of "conflation" (Stump 2017). A solution along these lines clearly improves on Anderson's theory, which addresses the question of weak but not strong generative capacity. However, having both ordered and unordered rule blocks, or rule blocks and conflation, provides for a rather baroque structure that appears to work around what I consider a design flaw of a rule block approach: being amorphous, PFM may look like the simpler model as far as derived structure is concerned, but this comes at the expense of an overly elaborate derivation structure. Thus the absence of morphological structure at the top-level is more than compensated by having several layers of structure in the cascade of rules of exponence and conflation rules, with intermediate representations at every level. The morphous inferential-realisational analysis that I shall present in Section 5, by contrast, invokes no structure at all beyond the assumption that exponents are segmentable, an assumption which is by the way implicitly made by the PFM rule system.

While at first sight, the move from ordinary extended exponence to exuberant exponence appeared as a mere quantitative difference, exuberance is actually a game-changer, inducing a qualitative difference when confronted with concrete formal theories: while incremental theories can indeed be discarded en bloc, the ability to account for exuberant exponence does not align with the distinction between lexical-realisational and inferential-realisational theories. As we have seen there are approaches of either type that can successfully analyse the data, as well as approaches that fail to do so. That means that the ability to capture exuberant exponence does not depend so much on the broad affiliation within the typology of morphological theories but rather on the specifics of the formal implementation.

4 Harris's word-based approach

Harris (2009) herself proposes a word-based analysis of Batsbi class marking, inspired inter alios by Blevins (2006), see Blevins (2016) for a more recent reference. Under a word-based perspective, speakers are assumed to store paradigms of high frequency words and establish analogical relationships between the cells of the paradigm. Such analogical relations are abstracted from full or partial paradigms, their application enabling speakers to form new word forms from already memorised ones. For instance, given a stored paradigm, word-to-word relations between paradigm cells can be abstracted out, like the one in (11):

(11) [Gender *n*] ~ [CM_n-X] ↔ [Gender *m*] ~ [CM_m-X]

According to her, such abstract relations, or the concrete instantiations thereof, to gender/number features and their corresponding surface exponents, make it possible to infer new forms from known forms, e.g. *yet:ŏ* 's/he pours milk' from *det:ŏ* 's/he pours tea' (recall that agreement is with the absolutive, which is the object of a transitive in this case).

For Batsbi, Harris (2009) assumes that lexical items and affixes each give rise to two basic schemata, one that features a class marker (**d**-LEX/**d**-AFF), and one that does not (LEX/AFF). Based on the lexical schemata, Harris suggests that basic verbs like ***d**-ek'-$\tilde{i}^{n}$* 'they fell' and *ot'-ŏ* 'they spread it' can be schematised as [**d**-LEX]$_V$ and [LEX]$_V$, respectively.

She then moves on to "first order" extensions, including transitive and intransitive markers and suggests two abstract schemata [V-**d**-AFF]$_V$ and [V-AFF]$_V$ the first of which is instantiated in the following sub-schemata (Table 3.4).

Table 3.4: Transitive/intransitive first order extensions

	Sub-schema	Example	Translation
a.	[[**d**-LEX]$_V$ -**d**-i-]$_V$	**d**-ek'-**d**-iy-en	'threw it off'
b.	[[**d**-LEX]$_V$ -**d**-al]$_V$	**y**-opx-**y**-al-in=e	'dressed and'
c.	[[LEX]$_V$ -**d**-i-]$_V$	tat:-**b**-iy-en	'pushed it'
d.	[[LEX]$_V$ -**d**-al-]$_V$	oc'-**v**-al-in-es	'I weighed'

In order to incorporate second order extensions such as the evidential I and the aorist evidential, Harris (2009) proposes even more complex sub-schemata, illustrated in Table 3.5.

Table 3.5: Second order schemata

	Sub-schema	Explanation
a.	[[**d**-LEX]$_V$ -**d**-anŏ]$_V$	evidential I of simple verb with preradical CM
b.	[[[**d**- LEX]$_V$ -**d**-i]$_V$ -**d**-anŏ]$_V$	evidential I of derived transitive with preradical CM
c.	[[**d**- LEX]$_V$ -inŏ]$_V$	aorist evidential of simple verb with preradical CM
d.	[[[**d**- LEX]$_V$ -**d**-i]$_V$ -inŏ]$_V$	aorist evidential of derived transitive with preradical CM

As indicated by Harris (2009), the sub-schemata in Table 3.5 are only a subset of the actual number of schemata. Factoring in only the stem and transitive/intransitive schemata, the number grows to 16. Once we factor in TAM markers (e.g. present, imperfective or aorist), we end up with a considerably greater number. The word-based approach therefore does not appear to be a very economical way of capturing the dependency of a class marker on the marker that licenses it. What is more, such a view will hardly scale up to the description of morphologically even more complex languages. Finally, a word-based view misses the utterly local nature of licensing involved with class marking.

It is rather clear what the basic intuitions are that Harris intends to capture with her (informal) analysis: to account for the dependent nature of gender markers (via schemata) and their uniform pattern of alternation (via analogy). It is far less clear though how the different abstractions of intermediate structures that she offers are to be interpreted in a word-based model. As a result, there are two basic readings of her analysis that I shall assume as plausible for the rest of this chapter: a purely word-based view, where intermediate abstractions are just abbreviatory devices (Harris 2009: 298), or a constructional view where such abstractions are meant to have some theoretical status. Depending on which of the two readings is correct, the current paper will make a different contribution: if the latter, it will provide a formal interpretation of Harris (2009), leading to a clearer understanding of what the different variables (depicted in bold face or small caps) are and how they can be interpreted in a generative grammar that makes use of typed feature logic. If, however, the former, word-based interpretation is more faithful, it will show in addition how a schema-based approach of Batsbi can be formalised in a rigorous fashion, without necessitating a fully holistic, word-based view.

In the next section, I shall therefore present an alternative analysis of Batsbi exuberant exponence, one which completely avoids unfolding the entire morphotactics into primary and secondary sub-schemata, but relies instead on a typed feature logic to give a formal interpretation to the basic combination of class-markers and the exponents that license their occurrence.

5 Analysis

In this section, I shall present Information-based Morphology, an inferential-realisational theory of inflection and show how the two basic analytical devices, inheritance and cross-classification in typed feature structures, are sufficient to provide an analysis of Batsbi exuberant exponence that captures simultaneously the dependent nature of class markers and the uniformity of their exponence. Furthermore, this analysis will highlight how Harris' original proposal, when understood in constructional rather than word-based terms, can be given a straightforward formal interpretation using the IbM framework.

5.1 Information-based Morphology

In this section[5], I shall present the basic architecture of Information-based Morphology (IbM, Crysmann & Bonami 2016; Crysmann 2017), an inferential-realisational theory of inflection (cf. Stump 2001) that is couched entirely within typed feature logic, as assumed in HPSG (Pollard & Sag 1987; 1994). In IbM, realisation rules embody partial generalisations over words, where each rule may pair *m* morphosyntactic properties with *n* morphs that serve to express them. IbM is a morphous theory (Crysmann & Bonami 2016), i.e. exponents are described as structured morphs, combining descriptions of shape (=phonology) and position class. As a consequence, individual rules can introduce multiple morphs, in different, even discontinuous positions. By means of multiple inheritance hierarchies of rule types, commonalities between rules are abstracted out: in essence, every piece of information can be underspecified, including shape, position, number of exponents, morphosyntactic properties, etc.

In contrast to other realisational theories, such as Paradigm Function Morphology (Stump 2001) or A-morphous Morphology (Anderson 1992), IbM does away with procedural concepts such as ordered rule blocks. Moreover, rules in IbM are non-recursive, reflecting the fact that inflectional paradigms in general constitute finite domains. Owing to the absence of rule blocks, IbM embraces a strong notion of Pāṇini's principle or the elsewhere condition (Kiparsky 1985) which is couched purely in terms of informational content (=subsumption) and therefore applies in a global fashion (Crysmann 2017), thereby including discontinuous bleeding (Noyer 1992).

5.1.1 Inflectional rules as partial abstraction over words

From the viewpoint of inflectional morphology, words can be regarded as associations between a phonological shape (PH) and a morphosyntactic property set (MS), the latter including, of course, information pertaining to lexeme identity. This correspondence can be described in a maximally holistic fashion, as shown in Figure 3.1. Throughout this section, I shall use German (circumfixal) passive/past participle (*ppp*) formation, as witnessed by *ge-setz-t* 'put', for illustration.

[5]This section has been largely reproduced from Crysmann & Bonami (2017). For an overview of alternative approaches to morphology within HPSG and constraint-based grammar, please see Bonami & Crysmann (2016).

One difference between the current version of IbM and previous ones is that we have now settled on considering MPH as a list rather than a set.

$$\begin{bmatrix} \textsc{ph} & \text{gesetzt} \\ \textsc{ms} & \left\{ \begin{bmatrix} \textsc{lid} & \textit{setzen} \end{bmatrix}, \begin{bmatrix} \textsc{tma} & \textit{ppp} \end{bmatrix} \right\} \end{bmatrix}$$

Figure 3.1: Holistic word-level association between form (PH) and function (MS)

Since words in inflectional languages typically consist of multiple segmentable parts, realisational models provide means to index position within a word: while in AM and PFM ordered rule blocks perform this function, IbM uses a list of morphs (MPH) in order to explicitly represent exponence. Having morphosyntactic properties and exponents represented as sets and lists, standard issues in inflectional morphology are straightforwardly captured at the level of rules: cumulative exponence corresponds to the expression of *m* properties by 1 morph, whereas extended (or multiple) exponence corresponds to 1 property being expressed by *n* morphs. Overlapping exponence finally represents the general case of *m* properties being realised by *n* exponents. Figure 3.2 illustrates the word-level $m : n$ correspondence of lexemic and inflectional properties to the multiple morphs that realise it. By means of simple underspecification, i.e. partial descriptions, one can easily abstract out realisation of the past participle property, arriving at a direct representation of circumfixal realisation.

Word:

$$\begin{bmatrix} \textsc{ph} & \text{gesetzt} \\ \textsc{mph} & \left\langle \begin{bmatrix} \textsc{ph} & \text{ge} \\ \textsc{pc} & -1 \end{bmatrix}, \begin{bmatrix} \textsc{ph} & \text{setz} \\ \textsc{pc} & 0 \end{bmatrix}, \begin{bmatrix} \textsc{ph} & \text{t} \\ \textsc{pc} & 1 \end{bmatrix} \right\rangle \\ \textsc{ms} & \left\{ \begin{bmatrix} \textsc{lid} & \textit{setzen} \end{bmatrix}, \begin{bmatrix} \textsc{tma} & \textit{ppp} \end{bmatrix} \right\} \end{bmatrix}$$

Abstraction of circumfixation ($1 : n$):

$$\begin{bmatrix} \textsc{mph} & \left\langle \begin{bmatrix} \textsc{ph} & \text{ge} \\ \textsc{pc} & -1 \end{bmatrix}, \begin{bmatrix} \textsc{ph} & \text{t} \\ \textsc{pc} & 1 \end{bmatrix}, \ldots \right\rangle \\ \textsc{ms} & \left\{ \begin{bmatrix} \textsc{tma} & \textit{ppp} \end{bmatrix}, \ldots \right\} \end{bmatrix}$$

Figure 3.2: Structured association of form (MPH) and function (MS)

Yet, a direct word-based description does not easily capture situations where the same association between form and content is used more than once in the same word, as is arguably the case for Swahili (Stump 1993; Crysmann & Bonami 2016; 2017) or, even more importantly for Batsbi (Harris 2009). By way of introducing a level of R(EALISATION) R(ULES), reuse of resources becomes possible. Rather than expressing the relation between form and function directly at the word level, IbM assumes that a word's description includes a specification of which rules license the realisation between form and content, as shown in Figure 3.3.

Realisation rules (members of set RR) pair a set of morphological properties to be expressed, the morphology under discussion (MUD) with a list of morphs that

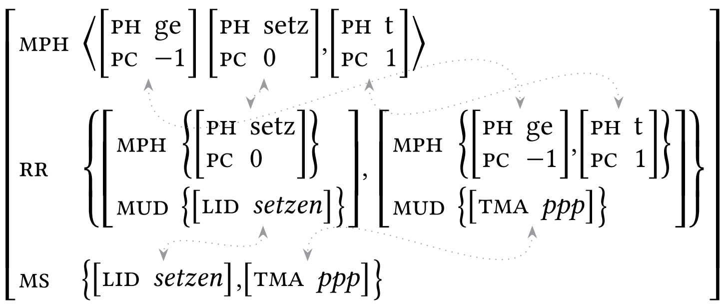

Figure 3.3: Association of form and function mediated by rule

realise them (MPH). In order to facilitate generalisations about shape and position in an independent fashion, IbM recognises each of them as first order properties of morphs, where PH represents a description of the phonological shape,[6] whereas PC corresponds to position class information. A general principle of morphological well-formedness (Figure 3.4) ensures that the properties expressed by rules add up to the word's property set and that the rules' MPH list add up to that of the word, i.e. no contribution of a rule may ever be lost.[7] In essence, a word's sequence of morphs, and hence, its phonology will be obtained by shuffling ($\bigcirc$) the rules' MPH lists in ascending order of position class (PC) indices (see Bonami & Crysmann 2013 for details). Similarly, a word's morphosyntactic property set (MS) will correspond to the non-trivial set union ($\uplus$) of the rules' MUD values.[8] Finally, the entire morphosyntactic property set of the word (MS) is exposed on each realisation rule by way of structure sharing ($\boxed{0}$).

This latter aspect, i.e. the relationship between MUD and MS in rule descriptions, surely deserves some more clarification in the context of this chapter. IbM makes a deliberate distinction between expression of a property and conditioning on a property: while MUD represents expression of properties, constraints on the MS set serve to capture allomorphic conditioning, in the sense of Carstairs (1987). There are two important consequences of this distinction (Crysmann 2017): first,

[6]For ease of presentation, I shall use strings to represent phonological contributions. More generally, PH(ON) value should be considered descriptions of phonological events, as suggested e.g. by Bird & Klein (1994).

[7]The principle of general well-formedness in Figure 3.4 bears some resemblance to LFG's principles of completeness and coherence (Bresnan 1982), as well as to the notion of "Total Accountability" proposed by Hockett (1947). Since $m : n$ relations are recognised at the most basic level, i.e. morphological rules, mappings between the contributions of the rules and the properties of the word can and should be $1 : 1$.

[8]While standard set union ($\cup$) allows for the situation that elements contributed by two sets may be collapsed, non-trivial set union ($\uplus$) insists that the sets to be unioned must be disjoint.

$$word \rightarrow \begin{bmatrix} \text{MPH} & \boxed{e_1} \circ \ldots \circ \boxed{e_n} \\ \text{RR} & \left\{ \begin{bmatrix} \text{MPH} & \boxed{e_1} \\ \text{MUD} & \boxed{m_1} \\ \text{MS} & \boxed{0} \end{bmatrix}, \ldots, \begin{bmatrix} \text{MPH} & \boxed{e_n} \\ \text{MUD} & \boxed{m_n} \\ \text{MS} & \boxed{0} \end{bmatrix} \right\} \\ \text{MS} & \boxed{m_1} \uplus \cdots \uplus \boxed{m_n} \end{bmatrix}$$

Figure 3.4: Morphological well-formedness

it becomes possible to make the application of inflectional rules a direct function of the information to be expressed, without having to postulate a system of (ordered) rule blocks. Second, it paves the way for a global notion of Pāṇinian competition, being able to distinguish between situations of discontinuous bleeding (Noyer 1992) and multiple or overlapping exponence. Thus, a rule with [MUD $\{ \alpha, \beta \}$] would preempt a rule with [MUD $\{ \beta \}$], since every morphosyntactic property is licensed (expressed) by exactly one rule. The rules [MUD $\{ \alpha \}$, MS $\{ \beta \}$] and [MUD $\{ \beta \}$], by contrast, would give rise to overlapping exponence (provided exponents do not compete for position). Here, expression of α is merely conditioned on a property that is independently expressed: β. See Crysmann (2017) for extensive discussion of preemption and overlapping exponence in Swahili.

Realisation rules conceived like this essentially constitute partial abstractions over words, stating that some collection of morphs jointly expresses a collection of morphosyntactic properties. In the example in Figure 3.3, we find that realisation rules thus conceived implement the $m : n$ nature of inflectional morphology at the most basic level: while the representation of classical morphemes as $1 : 1$ correspondences is permitted, it is but one option. The circumfixal rule for past participial inflection directly captures the $1 : n$ nature of extended exponence.

5.1.2 Levels of abstraction

The fact that IbM, in contrast to PFM or AM, recognises $m : n$ relations between form and function at the most basic level of organisation, i.e. realisation rules, means that morphological generalisations can be expressed in a single place, namely simply as abstractions over rules. Rules in IbM are represented as typed feature structures organised in an inheritance hierarchy, such that properties common to leaf types can be abstracted out into more general supertypes. This vertical abstraction is illustrated in Figure 3.5. Using again German past participles as an example, the commonalities that regular circumfixal *ge-...-t* (as in *gesetzt* 'put') shares with subregular *ge-...-en* (as in *geschrieben* 'written') can be generalised as the properties of a rule supertype from which the more specific leaves inherit. Note that essentially all information except choice of suffixal shape is associated with the supertype. This includes the shared morphotactics of the suffix.

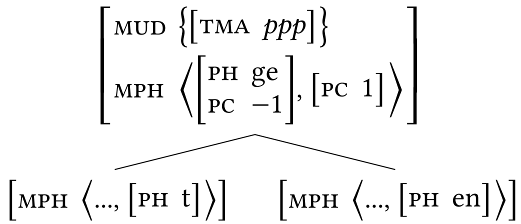

Figure 3.5: Vertical abstraction by inheritance

In addition to vertical abstraction by means of standard monotonic inheritance hierarchies, IbM draws on online type construction (Koenig & Jurafsky 1994): using dynamic cross-classification, leaf types from one dimension can be distributed over the leaf types of another dimension. This type of horizontal abstraction permits modelling of systematic alternations, as illustrated once more with German past participle formation:

(12) a. **ge**-setz-**t** ‘set/put’
b. über-setz-**t** ‘translated’
c. **ge**-schrieb-**en** ‘written’
d. über-schrieb-**en** ‘overwritten’

In the more complete set of past participle formations shown in (12), we find alternation not only between choice of suffix shape (*-t* vs. *-en*), but also between presence vs. absence of the prefixal part (*ge-*).

Figure 3.6 shows how online type construction enables us to generalise these patterns in a straightforward way: while the common supertype still captures properties true of all four different realisations, namely the property to be expressed and the fact that it involves at least a suffix, concrete prefixal and suffixal realisation patterns are segregated into dimensions of their own (indicated by PREF and SUFF). Systematic cross-classification (under unification) of types in PREF with those in SUFF yields the set of well-formed rule instances, e.g. distributing the left-hand rule type in PREF over the types in SUFF yields the rules for *ge-setz-t* and *ge-schrieb-en*, whereas distributing the right hand rule type in PREF gives us the rules for *über-setz-t* and *über-schrieb-en*, which are characterised by the absence of the participial prefix.

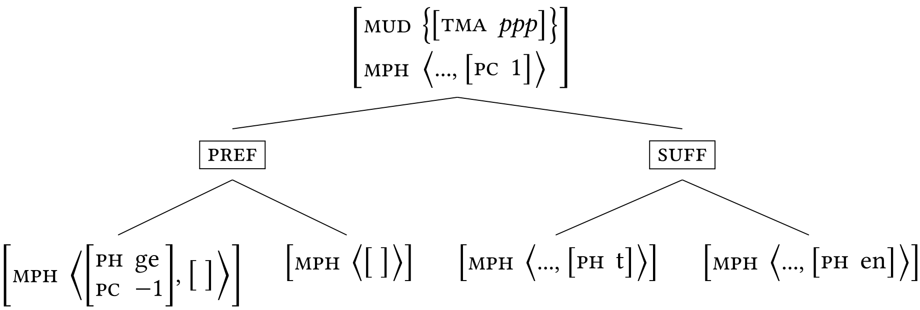

Figure 3.6: Horizontal abstraction by dynamic cross-classification

5.2 An information-based account of Batsbi exuberant exponence

Having introduced the basic workings of IbM, we are now in a position to approach the analysis of exuberant exponence in Batsbi. For the purposes of the following discussion, recall the two most central observations made in Section 2: first that the shape of class markers remains identical across all occurrences, and second, that the presence vs. absence of a class marker depends on their right-adjacent marker. Thus we saw both stems that trigger presence of an immediately preceding class marker, and stems that do not. Similarly, some classes of affixal exponents are obligatorily accompanied by a left-adjacent marker, whereas others do not license presence of such a marker. As a consequence, a word may surface with multiple identical class markers, a single pre-stem class marker or a single suffixal class marker, or even with no overt class marker at all.

The analysis I shall put forth in this section is that stems and affixes that trigger presence of overt agreement are actually allomorphically conditioned on gender marking properties, but that expression of gender marking can be zero, in the limiting case.

By way of illustration, let us start with a sample analysis of a word featuring exuberant exponence. As given in Figure 3.7, the word's MPH list features two occurrences of the gender V/VI plural marker *d*, each adjacent to a trigger, the stem *ek’* and the transitive marker *-iy*.

As indicated by coindexation, each instance of the agreement marker is introduced by the same realisation rule as its trigger, e.g. a single rule introduces both the stem *ek’* ($\boxed{b}$) and its dependent class marker ($\boxed{a}$). The same holds for the transitivity marker *-iy* ($\boxed{d}$) and its accompanying class marker ($\boxed{c}$). Each of these complex rules expresses some property other than class agreement, as indicated by their MUD value, e.g. lexemic identity ($\boxed{t}$), or transitivity ($\boxed{u}$), but both are

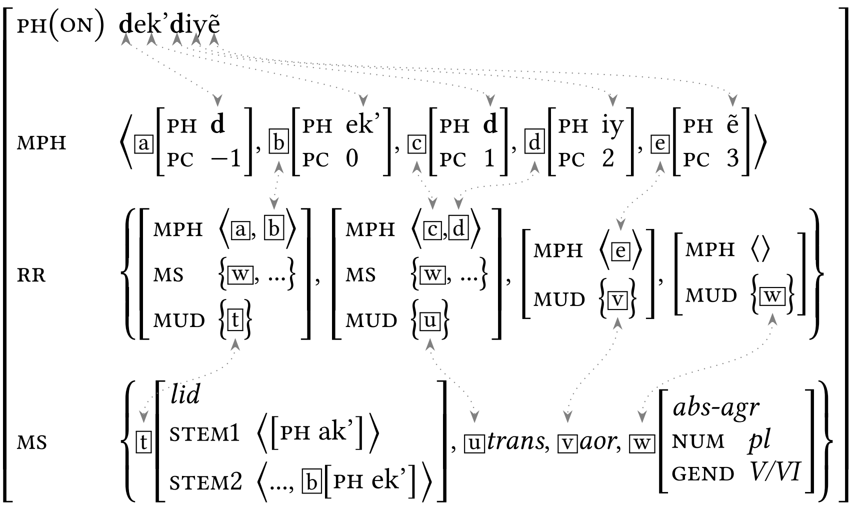

Figure 3.7: Sample analysis of Batsbi exuberant exponence

conditioned on the morphosyntactic property of gender/number agreement (w), specified as a constraint on the entire MS set. Since gender/number agreement has no expression independent of a trigger, and since in many words there is no overt exponent of class marking agreement, owing to the fact that only around 25% of stems and a select few suffixes license these dependent markers, I shall assume that class marking is expressed by default zero realisation, i.e. a rule that realises any property that has no more specific realisation rule by the empty set of morphs.[9] When class agreement does indeed surface, its dependent nature is best understood in terms of inflectional allomorphy.

5.3 Rule types for gender/number marking

Having sketched the overall line of analysis, I shall now present a description of the actual rule system starting with the type hierarchy that associates gender/number agreement features with any particular shape of class marker.

[9]This rule is similar in spirit to the identity function default of Stump (2001). Note that in IbM, just like in PFM, this kind of default reasoning is part of the logic, based entirely on the notion of information. Furthermore, it only applies between rule instances, i.e. leaves of the hierarchy, leaving multiple inheritance in the type hierarchy entirely monotonic. This contrasts sharply with Network Morphology (=NM Brown & Hippisley 2012), where defaults are used at the description level and at any node in the hierarchy, necessitating strong assumptions about orthogonality of properties in order to keep resolution of defaults sound. In the remainder of this chapter, I shall make no further reference to NM, for the simple reason that, as far as I am aware, the two areas under discussion here, i.e. multiple exponence and morphotactics, have not been the focus of research in that framework, making it difficult to assess its predictions.

At the top of the hierarchy in Figure 3.8, we find properties common to all class markers. First and foremost, the morphotactic description on MPH captures the fact that all class marking is dependent, consisting of two adjacent morphs. This basic property is expressed by means of requiring the list of morphs to be contributed by any class-marking rule to be bimorphic, i.e. a list of length 2. The phonology (PH) and position class (PC) of the morphs thus contributed are further constrained to have a consonantal morph immediately followed by a vowel-initial one, as dictated by the strictly consecutive position class indices. Second, the general rule type and its subtypes are restricted to have an *abs-agr* feature structure on the morphosyntactic property set.

Subtypes in the hierarchy in Figure 3.8 now further constrain the shape of the class marker. At the first level down in the hierarchy, the phonological shape of the initial consonantal marker is fixed. While *v-* is restricted to the singular of gender I and *d-* is treated as the default class marker, the two remaining markers *b-* and *j-* are both subject to unmotivated syncretism. This can be captured in a straightforward way by fixing their morphosyntactic constraints extensionally on the subtypes they dominate. This is possible since rule instances in IbM are only ever based on leaf types, following Koenig (1999).

As given in Figure 3.8 (page 74), the rule type for default CM marking is fully underspecified. The version of Pāṇini's principle that IbM assumes will actually preempt application of any more general rule in the presence of a more specific one.

(13) Pāṇinian competition (PAN) (Crysmann 2017)

1. For any leaf type t_1[MUD μ_1,MS σ], t_2[MUD μ_2,MS $\sigma \wedge \tau$] is a morphological competitor, iff $\mu_1 \subseteq \mu_2$.
2. For any leaf type t_1 with competitor t_2, expand t_1's MS σ with the negation of t_2's MS $\sigma \wedge \tau$: $\sigma \wedge \neg(\sigma \wedge \tau) \equiv \sigma \wedge \neg\tau$.

According to Pāṇinian competition, which is a closure operation on the type hierarchy, the MS set of the more general description for the default marker *d-* will end up being specialised to the description in Figure 3.9, which is essentially complementation with respect to the descriptions of its competitors.

5.4 Deconstructing class marking (suffixes)

Having introduced the partial constraints on the shape and position of the class markers, we are now in a position to bring them together with the suffixal markers on which they depend. The essential analytic device we shall rely on is online

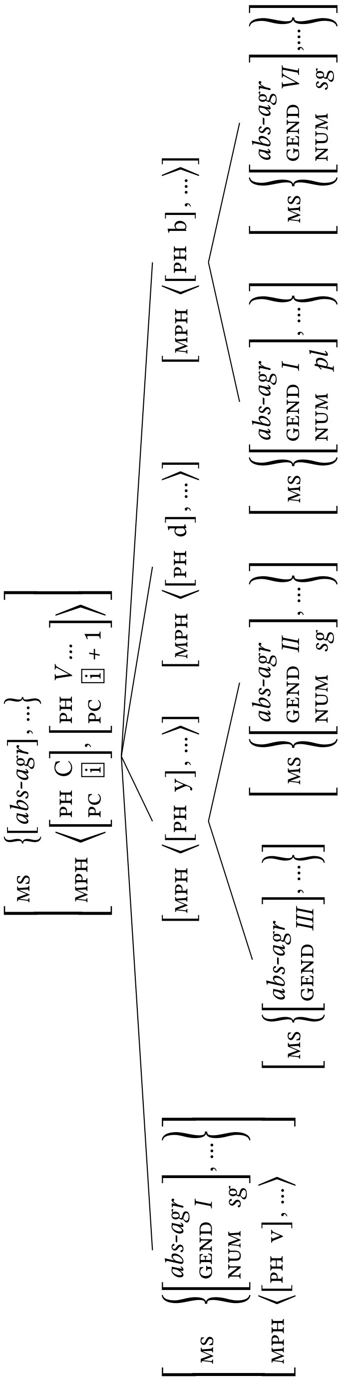

Figure 3.8: Subhierarchy of class marker rules

$$\begin{bmatrix} \text{MPH} & \left\langle \begin{bmatrix}\text{PH} & \text{d}\\ \text{PC} & \boxed{i}\end{bmatrix}, \begin{bmatrix}\text{PH} & \text{V...}\\ \text{PC} & \boxed{i}+1\end{bmatrix} \right\rangle \\ \text{MS} & \{abs\text{-}agr,...\} \wedge \begin{pmatrix} \neg\left\{\begin{bmatrix} abs\text{-}agr & \\ \text{GEND} & I\end{bmatrix},...\right\} \wedge \\ \neg\left\{\begin{bmatrix} abs\text{-}agr & \\ \text{GEND} & II \\ \text{NUM} & sg\end{bmatrix},...\right\} \wedge \\ \neg\left\{\begin{bmatrix} abs\text{-}agr & \\ \text{GEND} & III\end{bmatrix},...\right\} \wedge \\ \neg\left\{\begin{bmatrix} abs\text{-}agr & \\ \text{GEND} & VI \\ \text{NUM} & sg\end{bmatrix},...\right\} \end{pmatrix} \end{bmatrix}$$

Figure 3.9: Pāṇinian competition applied to default CM marker *d-*

type construction (Koenig & Jurafsky 1994), which enables us to state constraints on class markers and their licensors in dimensions of their own, yet distribute rule types in one dimension over the types in the other. Thus, each individual ingredient can be described in the most general way, while at the same time we can ensure their systematic combination.

The hierarchy of rule types in Figure 3.10 is organised into two dimensions, labelled ALLOMORPHY and EXPONENCE. In the former, one finds the type hierarchy of class marking from Figure 3.8, with class-marking leaf types abbreviated by the representative rule type for the *d*-marker. In the EXPONENCE dimension, we find realisation rule types for markers that show class-marking allomorphy, such as the present evidential or the intransitive, and some that do not. All realisation rules in this dimension specify a morphosyntactic property to be expressed via their non-empty MUD set, and all of them pair this property with a constraint on the exponent that serves to express this property, consisting of a phonological description and a position class index. The crucial difference between exponents that are accompanied by a class marker and those that are not is the constraint on the cardinality of the MPH set: while the latter specify a closed list (of length 1), those that do require a class marker are characterised by an open list.

Building on online type construction (Koenig & Jurafsky 1994), IbM obtains the set of rule instances by systematic intersection, under unification, of every leaf type from every dimension with every leaf type from every other dimension. The rule instances thus inferred from the type hierarchy are then subject to Pāṇinian competition.

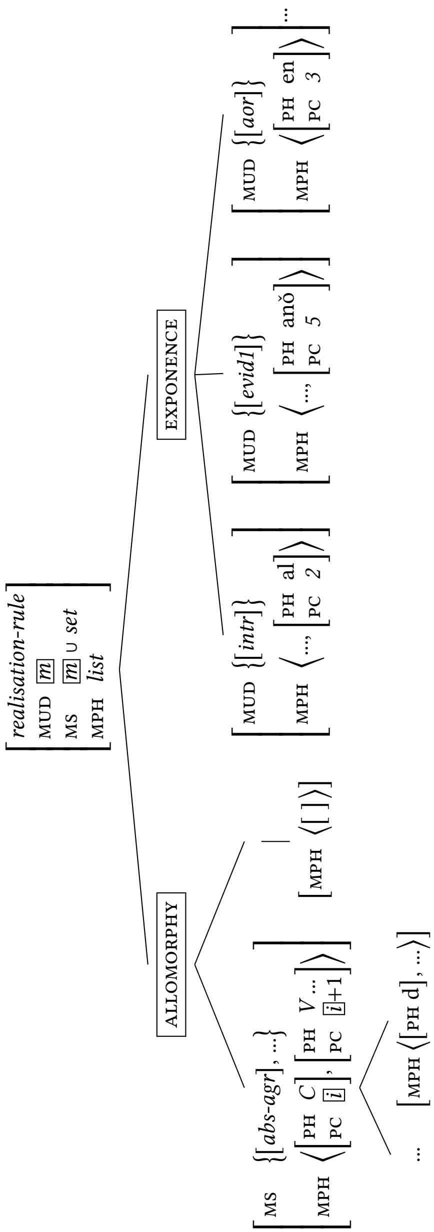

Figure 3.10: Hierarchy for suffix and class marking rule types

Rule types that do not take a class marker specify a monomorphic MPH set and therefore fail to unify with any of the class marking constraints, which are constrained to have a bimorphic MPH set, as specified on their supertype. Thus, rule types such as the one for the aorist can only combine with the rightmost leaf type in the ALLOMORPHY dimension, which merely constrains the cardinality of the MPH set to 1. Rule types that do take class markers, by contrast, do unify with the class marking constraints, yielding all combinations of class markers with the triggering marker.[10] When unifying class-marking and triggering types, unification of the phonological descriptions will ensure that morphs introduced in the two dimensions will receive the correct position class indices, thereby enforcing left adjacency of the class-marker to the triggering marker.

$$\begin{bmatrix} \text{MUD} & \left\{[\,]\right\} \\ \text{MPH} & \langle\rangle \end{bmatrix}$$

Figure 3.11: Default zero realisation

Finally, since expression of agreement properties does not necessarily have to be overt, I shall propose that agreement in Batsbi is expressed by a default rule of zero realisation, as shown in Figure 3.11. So any single morphosyntactic property that does not have any more specific expression can be realised without introducing any morphs. This will capture the vast number of cases where indeed no overt marking of agreement is found: as stated above, only a quarter of stems in the Batsbi lexicon license class agreement markers and only a select few affixes. If we assume that class agreement in Batsbi does not necessarily have an overt expression, we can treat those cases where we do find agreement as allomorphic variations of certain stems and affixes, as sketched in the analysis in Figure 3.7. Thus, by taking the majority case of zero exponence as our point of departure and treat dependent overt exponence as inflectional allomorphy, we avoid making arbitrary or even conflicting decisions about which overt exponents are realisations of agreement and which ones are just allomorphs.

5.5 Deconstructing class marking (stems)

As we have seen in Section 2.2, agreement marking of stems is ultimately decided in the lexicon: some stems take a class marker, some do not, and for some

[10]To be exact, triggering markers will also combine with the underspecified monomorphic rule type. However, these rules will always be preempted by the more specific rules showing allomorphic gender/number variation.

lexical entries we even find alternation where one stem in a lexeme's stem space comes with a class marker, but the other does not. To make sense of this lexically conditioned alternation, I shall build on the notion of stem spaces as proposed by Bonami & Boyé (2006). In IbM, stem spaces are provided by the lexeme and stem introduction rules, a subtype of realisation rules, serve to select an appropriate stem from the stem space and insert it into MPH (see Bonami & Crysmann (2018) for details on the interface between lexemes and the inflectional system).

As a first step to integrate inflecting and non-inflecting stems, I shall sketch a sample lexical entry for the alternating verb *ak'/ek'* and subsequently show how the general stem selection rules of the language will thread this lexemic information into the inflectional system, where it will take part in the allomorphic alternation we described above.

$$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle \text{NP}[\textit{abs}]_{\boxed{1}} \rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{RELS} & \left\langle \begin{bmatrix} \text{PRED} & \textit{fall} \\ \text{ARG1} & \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{MORPH} & \begin{bmatrix} \text{MS} & \left\{ \begin{bmatrix} \textit{lid} & \\ \text{STEM1} & \langle [\text{PH ak'}] \rangle \\ \text{STEM2} & \langle [\,], [\text{PH ek'}] \rangle \end{bmatrix}, \ldots \right\} \end{bmatrix} \end{bmatrix}$$

Figure 3.12: Sample lexical entry of a Batsi verb

At the lexical level, all it takes is to differentiate in the stem space between inflecting and non-inflecting stems. A most straightforward way of doing this is to replicate in the specification of stems a distinction we have already drawn for affixal markers, namely between monomorphic and bimorphic. Thus, an alternating stem such as *ak'/ek'* will have a singleton list as the value of STEM1, but a two-elementary list as the value of STEM2, as shown in Figure 3.12.

Stem introduction rules are given in the rule type hierarchy in Figure 3.13: just like the realisation rules for the aorist, evidential, transitive etc. in Figure 3.10 above, the stem introduction rules are part of the EXPONENCE dimension, so they are available for cross-classification with the class marking rule types. The two stem selection rules given here identify their MPH value with that of a stem value in MUD, STEM2 in the perfective, and STEM1 otherwise. Note that neither stem selection rule limits the arity of the stem values or of their MPH list. Thus, they both unify freely with any of the types in the ALLOMORPHY dimension, including all of the class-marking rule types, as well as the non-marking monomorphic type. Thus, cross-classification by online type construction will derive both bimorphic class-marking and monomorphic non-marking stem selection rules.

However, once any of the stem selection rules is applied to a concrete lexeme, bimorphic class marking rules will only be applicable to stem values of arity two, whereas monomorphic non-marking stem selection rules will exclusively apply to stem values of arity one.

To conclude, the present analysis of exuberant exponence in Batsbi exploits the fact that IbM recognises many-to-many relations between morphosyntactic properties at the most basic level of representation, namely realisation rules. Using online type construction in an inheritance hierarchy of rule types, the two most central generalisations regarding exuberant exponence in this language can be given a unified and straightforward account, by separating constraints on the shape of class markers from licensing their presence: Thus, while triggering affix rules and stems ultimately decide on whether they must (or may not) combine with a class marker, the constraints on class-marking are stated separately, distributing over rules of exponence.

5.6 Reflections on the dependent nature of exuberant exponence

The kind of exuberant exponence expounded in Batsbi witnesses two important properties: first, agreement marking is dependent on an adjacent triggering marker, a stem or some affix, and the number of class markers found then depends on the number of triggering stems or suffixes present in the word, yielding a variable degree of exuberant exponence. The formal analysis does justice to these two observations by treating the dependent class marker as morphologically conditioned allomorphy of the triggering stem or suffix. This raises the obvious question whether exuberant exponence must in general be of the dependent type.[11] Fully redundant multiple exponence involving more than two markers is rare, so I shall extrapolate from what we know about multiple exponence in general.

To answer this question, let us consider pre-prefixation in Nyanja (Stump 2001; Crysmann 2017): in this language a subclass of adjectives takes two agreement markers, one from the set of adjectival markers, the other from the set of verbal agreement markers.

(14) a. ci-lombo ci-kula.
CL7-weed CONC7-grow
'A weed grows.'

b. ci-manga ca-bwino
CL7-maize QUAL7-good
'good maize'

[11] Thanks to Jean-Pierre Koenig for drawing my attention to this.

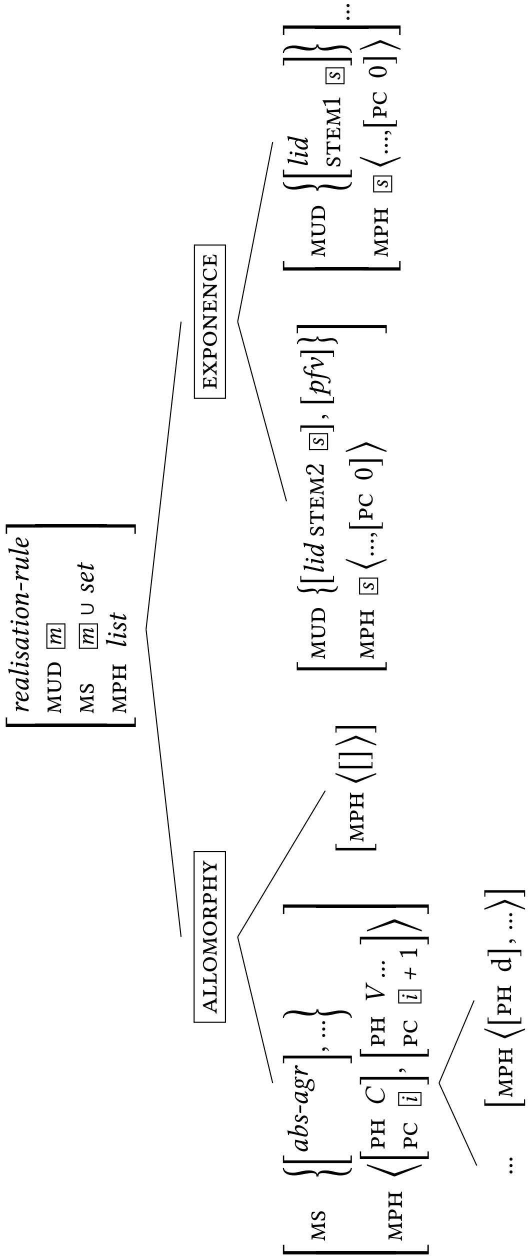

Figure 3.13: Stem selection and class-marking rule types

(15) ci-pewa ca-ci-kulu
CL7-hat(7/8) QUAL7-CONC7-large
'a large hat'

Multiple exponence in Nyanja is solely determined by inflection class membership, and the two agreement markers surface adjacent to each other, without any additional triggering exponent. In IbM, this situation has been analysed by means of composing simple verbal and adjectival markers into a class-specific morphotactically complex marker (Crysmann 2017). However, what we find here is composition of similar yet non-identical markers, each of which is attested independently.

The crucial difference between Nyanja and Batsbi is that the number of exponents is fixed in the former for any given inflection class, but it is variable and dependent on the presence of concrete stems and suffixes in the latter. Whenever multiple exponence is morphotactically dependent, the formal approach sketched here, which composes each instance of multiple marking with a triggering exponent, is to be preferred. It so happens that this approach is also much more apt at handling variable degrees of exuberance, a property that is actually expected, if exponence is dependent on a triggering marker. Composition among the instances of multiple exponence, by contrast, is the way to go, if multiple exponence is morphotactically independent and fixed with respect to the degree of exuberance.

6 Conclusion

In this paper I have discussed exuberant exponence in Batsbi (Harris 2009). I have shown that the design property of IbM to recognise $m : n$ relations between form and function at the level of realisation rules lends itself naturally to accounting for the dependent nature of these markers. Thus, under the perspective offered here, exuberant class marking in Batsbi is just a case of allomorphy on the markers/stems they depend on, conditioned by number and gender properties. The uniformity of shape of these markers has been captured by a system of cross-classifying type hierarchies along the dimensions of allomorphy and exponence, building on the formal notion of online type construction (Koenig & Jurafsky 1994) standardly embraced by IbM. As a result, I have offered a theory of Batsbi exuberant exponence that is as holistic as necessary to capture dependence, and at the same time as atomistic as possible, thereby facilitating reuse. In other words, the current approach captures the constructional properties of the system within a formal generative model.

Finally, this paper provided some meta-theoretical result, showing that there is only limited a priori superiority of inferential-realisational approaches over lexical-realisational ones: just as much as the conceptual foundations, it is the formal expressivity of the specific framework that determines its adequacy in light of exuberant exponence.

Abbreviations

The glosses in this chapter follow the original description by Harris (2009), slightly adapted to adhere more fully to the Leipzig conventions. Here is a list of additional abbreviations being used: CM (class marker), PRES (present), AOR (aorist), EVID1 (evidential 1), CON (contact case). Furthermore, inherent noun class is indicated by means of the exponents of the singular and plural class markers.

Acknowledgements

Preliminary versions of this work have been presented at the 16th international morphology meeting (Budapest, May 2014), the 3rd European workshop on Head-driven Phrase Structure Grammar (Paris, November 2014), and the workshop on building blocks (Leipzig, November 2014). I would like to thank the audiences at these venues for their questions and comments, in particular, Doug Arnold, Matthew Baerman, Olivier Bonami, Bernard Fradin, Alain Kihm, Jean-Pierre Koenig, Frank Richter, Barbara Stiebels, Greg Stump and Geraldine Walther.

This work was partially supported by a public grant overseen by the French National Research Agency (ANR) as part of the program "Investissements d'Avenir" (reference: ANR-10-LABX-0083). It contributes to the IdEx Université de Paris (ANR-18-IDEX-0001).

References

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Bird, Steven & Ewan Klein. 1994. Phonological analysis in typed feature systems. *Computational Linguistics* 20(3). 455–491.

Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 42(3). 531–573. DOI: 10.1017/S0022226706004191.

Blevins, James P. 2016. *Word and paradigm morphology*. Oxford: Oxford University Press.

Bonami, Olivier & Gilles Boyé. 2006. Deriving inflectional irregularity. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on HPSG*, 39–59. Stanford: CSLI Publications.

Bonami, Olivier & Berthold Crysmann. 2013. Morphotactics in an information-based model of realisational morphology. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 27–47. Stanford: CSLI Publications.

Bonami, Olivier & Berthold Crysmann. 2016. The role of morphology in constraint-based lexicalist grammars. In Andrew Hippisley & Gregory T. Stump (eds.), *Cambridge handbook of morphology*, 609–656. Cambridge: Cambridge University Press.

Bonami, Olivier & Berthold Crysmann. 2018. Lexeme and flexeme in a formal theory of grammar. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer (eds.), *The lexeme in descriptive and theoretical morphology* (Empirically Oriented Theoretical Morphology and Syntax), 175–202. Berlin: Language Science Press. DOI: 10.5281/zenodo.1402520.

Bresnan, Joan (ed.). 1982. *The Mental Representation of Grammatical Relations*. Cambridge, Massachusetts: The MIT Press.

Brown, Dunstan & Andrew Hippisley. 2012. *Network Morphology: A defaults-based theory of word structure* (Cambridge Studies in Linguistics). Cambridge University Press. DOI: 10.1017/CBO9780511794346.

Carstairs, Andrew. 1987. *Allomorphy in inflection*. London: Croom Helm.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Crysmann, Berthold. 2017. Inferential-realisational morphology without rule blocks: An information-based approach. In Nikolas Gisborne & Andrew Hippisley (eds.), *Defaults in morphological theory*, 182–213. Oxford: Oxford University Press.

Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics in Information-based Morphology. *Journal of Linguistics* 52(2). 311–374. DOI: 10.1017/S0022226715000018.

Crysmann, Berthold & Olivier Bonami. 2017. Atomistic and holistic exponence in Information-based Morphology. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, Lexington, KY*, 141–161. Stanford: CSLI Publications.

Halle, Morris & Alex Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from Building 20*, 111–176. Cambridge: MIT Press.

Harris, Alice C. 2009. Exuberant exponence in Batsbi. *Natural Language and Linguistic Theory* 27. 267–303. DOI: 10.1007/s11049-009-9070-8.

Harris, Alice C. 2017. *Multiple exponence*. Oxford: Oxford University Press.
Hockett, Charles F. 1947. Problems of morphemic analysis. *Language* 23. 321–343.
Holisky, Dee Ann & Rusudan Gagua. 1994. Tsova-Tush (Batsbi). In Rieks Smeets (ed.), *The indigenous languages of the Caucasus 4: North East Caucasian, part 2*, 147–212. Delmar: Caravan Press.
Kadagiʒe, Davit & Niko Kadagiʒe. 1984. *C'ova-tušur-kartul-rusuli leksik'oni*. Tbilisi: Mecniereba.
Kiparsky, Paul. 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2. 83–136.
Koenig, Jean-Pierre. 1999. *Lexical relations*. Stanford: CSLI Publications.
Koenig, Jean-Pierre & Daniel Jurafsky. 1994. Type underspecification and on-line type construction. In *Proceedings of WCCFL XIII*, 270–285.
Matthews, Peter H. 1972. *Inflectional morphology. A theoretical study based on aspects of Latin verb conjugation*. Cambridge: Cambridge University Press.
Müller, Gereon. 2007. Extended exponence by enrichment: Argument encoding in German, Archi, and Timucua. *University of Pennsylvania Working Papers in Linguistics* 13(1).
Noyer, Rolf. 1992. *Features, positions and affixes in autonomous morphological structure*. MIT. (Doctoral dissertation).
Pollard, Carl & Ivan A. Sag. 1987. *Information-based syntax and semantics*. Stanford: CSLI Publications.
Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Stanford: CSLI Publications; The University of Chicago Press.
Steele, Susan. 1995. Towards a theory of morphological information. *Language* 71(2). 260–309.
Stump, Gregory T. 1993. Position classes and morphological theory. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of morphology 1992*, 129–180. Dordrecht: Kluwer.
Stump, Gregory T. 2001. *Inflectional morphology. A theory of paradigm structure* (Cambridge Studies in Linguistics 93). Cambridge: Cambridge University Press.
Stump, Gregory T. 2017. Rule conflation in an inferential-realizational theory of morphotactics. *Acta Linguistica Academica* 64(1). 79–124.
Wunderlich, Dieter & Ray Fabri. 1995. Minimalist Morphology: an approach to inflection. *Zeitschrift für Sprachwissenschaft* 14(2). 236–294.