

Against Split Morphology: 'inflection', 'derivation', and the structure of the SBCG lexicon

Robert Hepburn-Gray
University at Buffalo (SUNY)
rhepbur@buffalo.edu

1 Assumptions of (Split-)Morphological Theory

- ▶ The following propositions are often assumed in morphological theory
 - a) Inflection/derivation is a robust, universal distinction (or at least 'the norm'). Distinction can be more fine grained (Sag), but not less?
 - b) There exists a lexeme/word distinction, supported by an 'intuitive' (Booij, Stump?) difference between pairs like tax/taxes and tax/taxation
 - c) Metatheoretical assumption: typologically diverse languages should be analyzed using a single (though general) morphological framework
- ▶ Assumption A is referred to by Bauer (1997) as SPLIT MORPHOLOGY.
- ▶ Criteria for distinguishing inflection from derivation according to Stump (2005, pp. 53-58):
 - a) Derivation can change part-of-speech class, while inflection cannot
 - b) Inflection applies to a category without exception; derivation applies sporadically
 - c) Inflection is semantically regular; derivation is frequently less than fully semantically regular
 - d) Inflection is syntactically determined; derivation is not
 - e) Derivational processes apply before inflectional processes
- ▶ Stump finds exceptions to each of these criteria, but retains the distinction.
- ▶ Assumption B is intimately connected to Assumption A, with inflection creating words from lexemes, and derivation creating new lexemes from old ones.

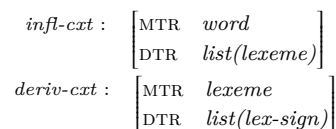


Figure 1: Split Morphology in SBCG (Sag 2012)

- ▶ Note however, Stump (2001, p. 253): “[B]oth inflectional paradigms and derivational paradigms are inventories projected from a single lexeme.”

- ▶ Data presented here from Niger-Congo noun class systems suggests that these systems do not display an inflection/derivation distinction.
- ▶ As one of the world’s largest language families (approx. 1500 in Ethnologue), expectations to the split-morphology type should not be considered marginal, and categories of morphological construction should not be assumed when analysing newly documented languages.
- ▶ Moreover, the diversity of morphological constructions found throughout the world’s languages suggests that fundamentally different analyses may be required for different constructions (perhaps even within the same language).

2 Typology of Niger-Congo Noun Class Systems

- ▶ Properties typical of Niger-Congo noun class systems (Kießling 2013, pp. 44-45)
 - a) all nouns assigned to a limited set of noun classes
 - b) all nouns control, by virtue of their assignment to a class, a system of concordial agreement which penetrates vast sections of the morphosyntax
 - c) class assignment is governed by semantic principles so that classes could be described as semantic networks (but not necessarily synchronically active/cognitively real (Dingemanse 2006, pp. 22-23))
 - d) most noun classes form singular-plural pairs or genders
- ▶ Otoro, Kordofanian (Stevenson 2009)

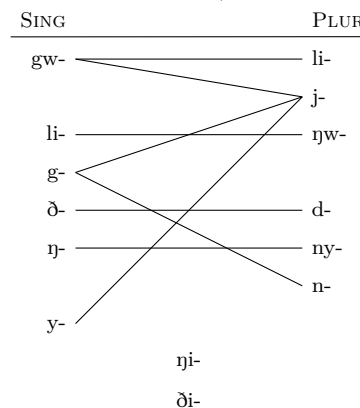


Table 1: Example Otoro Paradigms

GEN	SING	PLUR	GLOSS
<i>gw-/li-</i>	gwiji	liji	'person'
<i>gw-/j-</i>	gwaɽe	jaɽe	'tree'
<i>g-/j-</i>	gilöð	jilöð	'hoe'
<i>ð-/j-</i>	ðimu	jimu	'scorpion'

Figure 2: Otoro NC System

- ▶ As is often the case in Niger-Congo noun class systems, there are classes which

participate in multiple ‘genders’, such as gw-, j-, and g-, which form pairs with multiple classes, and ð- and y-, which participate in single and double class genders

- ▶ Moreover, number is present semantically, but is not an active morphosyntactic feature (Welmers 1973)
- ▶ In Indo European, there exist patterns (such as SV agreement) which are sensitive to number, but not gender. In Niger-Congo, however, one often finds systems where there exist no constructions which are sensitive to number distinct from class.

3 Number as a derivational process?

- ▶ Lumun (Smits 2011) represents a particularly irregular number system.

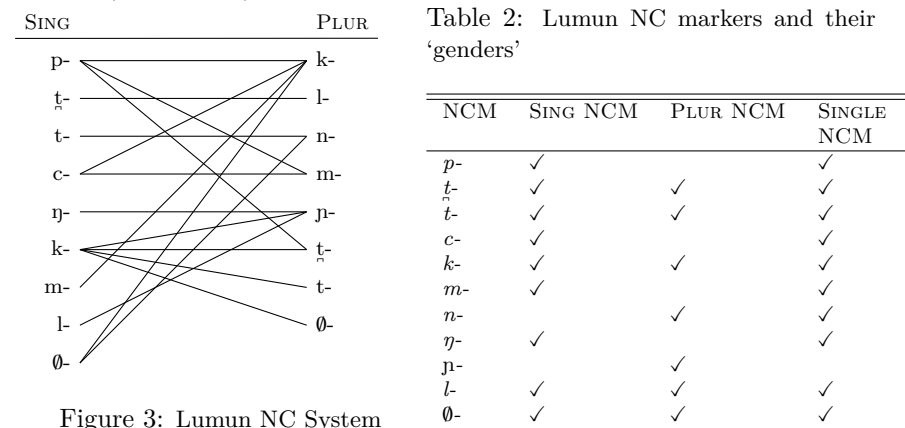


Figure 3: Lumun NC System

Table 2: Lumun NC markers and their ‘genders’

NCM	SING NCM	PLUR NCM	SINGLE NCM
<i>p-</i>	✓		✓
<i>t̄-</i>	✓		✓
<i>t-</i>	✓	✓	✓
<i>c-</i>	✓		✓
<i>k-</i>	✓	✓	✓
<i>m-</i>	✓		✓
<i>n-</i>		✓	✓
<i>ŋ-</i>	✓		✓
<i>j̄-</i>		✓	✓
<i>l-</i>	✓	✓	✓
<i>∅-</i>	✓	✓	✓

Table 3: The nine ‘genders’ of class marker *k-*

GEN	SING	PLUR	GLOSS
<i>k-/∅-</i>	kùmmùk	ùmmùk	‘pot/pots’
<i>k-/t̄-</i>	kupú	t̄upú	‘peice of k.o wood/k.o wood’
<i>k-/t-</i>	kua	tua	‘strand of hair/hair’
<i>k-/j̄-</i>	kùkkú	j̄ùkkú	‘groundnut/ groundnuts’
<i>p-/k-</i>	pura	kira	‘tree/forest’
<i>c-/k-</i>	cít	kít	‘eye/eyes’
<i>∅-/k-</i>	iké	kiké	‘giraffe/giraffes’
<i>k-</i>	kàrèt		‘abusive language’

- ▶ 26 distinct genders from only 11 different class markers
- ▶ Number is NOT semantically regular, with functions such as collective and

singulative as well

- ▶ Smits argues, following Schadeberg (2001), that number marking should be considered a derivational rather than inflectional process.
- ▶ Problems:
 - a) Gender-based analysis would posit widespread accidental homophony
 - b) Agreement is inflection par excellence. An analysis where class marking on nouns is derivational but inflectional on agreement targets is highly undesirable.

4 Paradigm Networks

- ▶ Paradigm networks such as the following can be found throughout the Niger-Congo family (Hepburn-Gray 2016).

Table 4: Botanical Paradigm Network in Bainounk (Cobbinah 2013, p. 319)

NC Paradigm	- <i>dooma</i> ‘kaba’	- <i>taat</i> ‘annona’
<i>si-/mun-</i>	‘kaba tree’	‘annona tree’
<i>bu-/i-/di-</i>	‘kaba fruit’	‘annona fruit’
<i>ja-</i>	‘leaves of the kaba tree’	‘leaves of the annona tree’

Table 5: Ethnic Group Paradigm in Cicipu (McGill 2007, p. 61)

Class	Acipu	Karishen	Kadonho	Hausa	Gloss
8	c-cípù	∅-rìsìnò	d-dípó	k-kógó	Person
2	à-cípù	ò-rìsìnò	ò-dípó	ò-kógó	People
1		kò-rìsìnò	kò-dípó		Town/Area
6	ci-cípù	tì-rìsìnò	tì-dípó	tì-kógó	Language

- ▶ Problems:
 - a) no principled way to choose a base lexeme, from which others are derived
 - b) could posit separate lexeme from which all are derived, but this lexeme never surfaces then
- ▶ Koenig (1999, p. 150) dicusses a similar example in English: *regress/regressive/regression* vs. **agress/agressive/agression*. The absence of the verb *agress* is explained as a missing root, which is only constructionally introduced in the *agressive/agression* constructions. However, it is not a root that is missing, but the verb cell in a derivational paradigm.

5 Paradigms in Morphological Theory

- ▶ Stump and Finkel (2013) distinguish between the canonical extremes of the PURE WORD-AND-PARADIGM MORPHOLOGY (PWPM) hypothesis and the PURE EXPONENCE-BASED MORPHOLOGY (PEM) hypothesis.

Table 6: Differences between the PWPM and PEM hypotheses (Stump and Finkel 2013, p. 265)

CRITERION	PWPM	PEM
IC membership	represented by means of a set of lexically listed principal parts	represented by means of a diacritic+one or more stems
Rules	implicative rules formulated in terms of realized cells	rules of exponence formulated in terms of stems

- ▶ S&F Argue that the principal part/diacritic distinction is a false one, but use diacritic in formalism
- ▶ For rules, they propose hybrid model with rules of exponence as primary and implicative rules of referral to cover syncretism
- ▶ Evidence for necessity of rules of exponence based on stem variation in Sanskrit. Since diversity is assumed here, rules of exponence are not motivated for these Niger-Congo languages.
- ▶ The efficiency of paradigm diacritics for languages with large inflectional paradigms is less obvious for these NC languages, since the paradigm diacritic only specifies two paradigm cells, both of which can be specified already by a single morphosyntactic feature (CLASS).
- ▶ Introduction of paradigm diacritics would be akin to introducing the notion of gender to these languages. See Schadeberg (2001) for arguments against a 'gender-based' analysis of Swahili.

6 Formalism

- ▶ Propose a WORD AND PARADIGM (Blevins 2006) model of morphology for these Niger-Congo languages.
- ▶ In the theory proposed here, a lexeme is not a type of sign. A lexeme is simply the knowledge that a set of words is paradigmatically related, and the information that is shared between this set of words. (See Blevins (2006) for the notion of 'abstractive' stems, lexemes etc.)
- ▶ Paradigm cells which are not principle parts are generated via analogy (*an-ctx*).

Compounds (*comp-ctx*) remain a distinct construction type, as somewhat of a hybrid word/phrase construction.

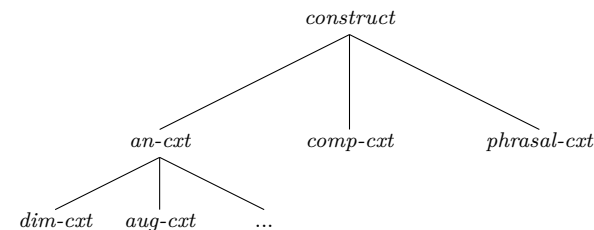


Figure 4: Construct types

- ▶ I adapt the following from Koenig (1999).

$$head \Rightarrow \left[\begin{array}{ll} \text{LXM} & lex-prop \\ \mu\text{-FEAT(URES)} & \mu-prop \end{array} \right]$$

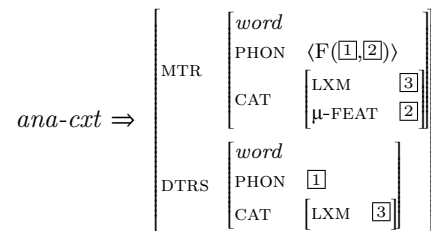
Figure 5: Head Feature Types Signature

- ▶ The following type signature for the sign type *word* encodes the semantics associated with the lexeme.

$$\left[\begin{array}{l} word \\ \text{MORSYN | HEAD} \\ \text{SEM} \end{array} \left[\begin{array}{l} \text{LXM} \\ \mu\text{-FEAT} \\ \text{FRAMES} \end{array} \left[\begin{array}{ll} \text{LBL} & lexm-lbl \\ \text{FRAME} & \mathbb{1} \\ \text{CLASS} & class \end{array} \right] \right] \right]$$

Figure 6: Type Signature of Sign type *word*

- ▶ Here L represents the (possibly empty) set of semantic frames which may be associated with a particular paradigm cell (past time reference, etc.).
- ▶ Different types of *an-ctx* take the principle part of a lexeme, and associate the morphosyntac features of the desired paradigm cell with the new word.
- ▶ The phonology of the word is determined by a function, which takes as input the phonology of the principle part and the features of the new paradigm cell. This function contains the 'implicative rules' of the PWPM approach, which generates a proportional analogy with the corresponding exemplar paradigm.

Figure 7: Type Signature of *an-cxt*

► An example of an analogical construction

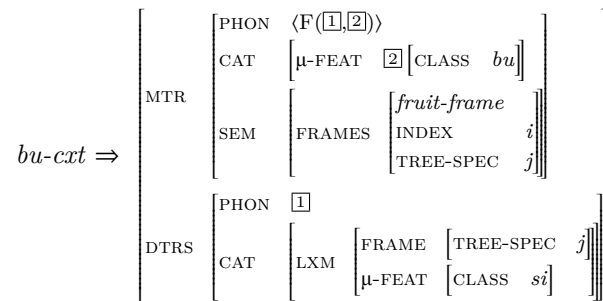


Figure 8: The 'fruit' construction of the botanical paradigm network

References

- Bauer, Laurie (1997). "Derivational Paradigms". In: *Yearbook of Morphology 1996*. Ed. by Jaap van Marle and Geert Booij. Dordrecht: Kluwer, pp. 243–256.
- Blevins, James (2006). "Word-based Morphology". In: *Journal of Linguistics* 42, pp. 531–573.
- Cobbinah, A.Y. (2013). "Noun classification and verbal nouns in Bãinounk Gubëeher". dissertation. London: University of London.
- Dingemanse, M. (2006). "The semantics of Bantu noun classification". M.A. Leiden: Leiden University.
- Hepburn-Gray, Robert (2016). "A Survey of Niger-Congo Noun Class Agreement Systems". In: *Towards Proto-Niger-Congo: Comparison and Reconstruction (2nd International Congress)*. Paris.
- Kießling, Roland (2013). "On the origin of Niger-Cong noun classification". In: *Historical Linguistics 2011: Selected papers from the twentieth International*

- Conference on Historical Linguistics, Osaka, 25-30 July 2011*. Amsterdam: John Benjamins, pp. 43–65.
- Koenig, Jean-Pierre (1999). *Lexical Relations*. Stanford: CSLI Publications.
- McGill, Stewart (2007). "The Cicipu noun class system". In: *Journal of West African Languages* 34.2, pp. 51–90.
- Sag, Ivan (2012). "Sign-Based Construction Grammar: An Informal Synopsis". In: *Sign-Based Construction Grammar*. Ed. by Ivan Sag and Hans Boas. Stanford: CSLI Publications, pp. 61–187.
- Schadeberg, Thilo (2001). "Number in Swahili grammar". In: *Afrikanistische Arbeitspapiere* 68, pp. 7–16.
- Smits, Heleen (2011). "Lunun noun classes and number". In: *Afrikanische Sprachen im Fokus. Linguistische Beiträge zum 19. Afrikanistentag, Mainz, 8.-10. April 2010*. Cologne: Rudiger Koppe Verlag, pp. 271–283.
- Stevenson, Roland C. (2009). "A Grammar of Otoro". In: *Tira and Otoro: Two Grammars of Kordofanian Languages by Roland C. Stevenson*. Ed. by Thilo Schadeberg. Cologne: Rüdiger Köppe Verlag.
- Stump, Gregory (2001). *Inflectional Morphology*. Cambridge: Cambridge University Press.
- (2005). "Word Formation and Inflectional Morphology". In: *Handbook of Word Formation*. Dordrecht: Springer, pp. 49–71.
- Stump, Gregory and Raphael A. Finkel (2013). *Morphological Typology: From Word to Paradigm*. Cambridge: Cambridge University Press.
- Welmers, William E. (1973). *African language structures*. Berkeley: UC press.