

---

**Good things come in small languages:  
grammatical loss and innovation in Nzadi**

Larry M. Hyman

---

Proceedings of Conference on  
**Language Documentation & Linguistic Theory 2**

Edited by Peter K. Austin, Oliver Bond, Monik Charette,  
David Nathan & Peter Sells

13-14 November 2009 School of Oriental and African Studies, University of London

Hans Rausing Endangered Languages Project  
Department of Linguistics  
School of Oriental and African Studies  
Thornhaugh Street, Russell Square  
London WC1H 0XG  
United Kingdom

Department of Linguistics:  
Tel: +44-20-7898-4640  
Fax: +44-20-7898-4679  
linguistics@soas.ac.uk  
<http://www.soas.ac.uk/academics/departments/linguistics>

Hans Rausing Endangered Languages Project:  
Tel: +44-20-7898-4578  
Fax: +44-20-7898-4349  
elap@soas.ac.uk  
<http://www.hrelp.org>

© 2009 Larry M. Hyman

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, on any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the author(s) of that part of the publication, except as permitted by UK copyright law.

ISBN: 978-0-7286-0392-9

This publication can be cited as:

Larry M. Hyman. 2009. Good things come in small languages: grammatical loss and innovation in Nzadi. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS.

or:

Larry M. Hyman. 2009. Good things come in small languages: grammatical loss and innovation in Nzadi. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS. [www.hrelp.org/eprints/ldlt2\\_02.pdf](http://www.hrelp.org/eprints/ldlt2_02.pdf)

# **Good things come in small languages: grammatical loss and innovation in Nzadi**

LARRY M. HYMAN

*University of California, Berkeley*

## 1. INTRODUCTION

In this paper I have two goals. First, given the nature of this conference, I would like to provide a few personal observations on the relation between language documentation and linguistic theory, particularly as it has concerned my work on African languages. Second, as an illustration of this work, I present some rather puzzling discoveries from Nzadi, a previously undocumented Bantu language spoken in the Democratic Republic of the Congo, which we were able to study in a field methods course, and on which Thera Crane, Simon Nsielanga Tukumu and I are preparing a (modest) descriptive grammar. What follows is a sketch of the paper I will present at the conference (with the conclusion still in progress).

## 2. LANGUAGE DOCUMENTATION AND LINGUISTIC THEORY

As I have told the organizers, I am deeply honored to have been invited to be a plenary speaker at LDLT2. On behalf of linguists who work as I do, I find it reassuring to be included among those whose interests combine both LD and LT, especially as I may not meet one's prototype of either: On the one hand, aside from field trips totalling two years in Africa some time ago, the 'documentation' that I currently do takes place *ex situ*, often with a single speaker. Except for the recording of a few narratives, data collection consists largely of elicitation. On the other hand, for many, 'linguistic theory' necessarily implies a preoccupation with the development and implementation of specific formal models. While much of my work has been theoretical in this sense, my research projects have been largely driven by the descriptive, historical and/or typological issues that arise in the languages I study - which have sometimes propelled me in unexpected directions (as will become apparent in section 3). Concerning my main subfield, phonology, modeling seems to be going in all directions. While many phonologists currently find solace in the concrete, phonetic grounding of phonological facts, others continue to derive satisfaction from highly abstract modeling, perhaps even what Prince (2007) refers to as 'free standing theories'. Although one can become be guided by either, the goal of phonology isn't to measure OR to formalize: it is to gain insight. Questions which typically arise are: What is going on here? Why is it the way it is? What can we relate it to? What draws many people to the study of languages is the potential to combine the excitement of discovery, however made, and insight, however expressed. To be successful in answering such questions, it seems intuitive that documentation and theory go ideally hand in hand.

As I have previously suggested concerning the notion of linguistic fieldwork (Hyman 2001), it would be reassuring if the field could adopt a broad interpretation of both language documentation and linguistic theory. With this goal in mind, let me present a formula suggested by my colleague Andrew Garrett (personal communication):

- (1) Field linguistics = Documentation + Description + Archiving

As indicated, there are three components to what Garrett identifies as field linguistics, which he goes on to define as follows:

- (2) Documentation is the activity of collecting data that might be used in a description (or might be used by somebody publishing a text collection, or making a cd of traditional songs).

*Description* is the activity of writing up statements about how the grammar works, basically, everything that might go into a useful grammar or part of a grammar or grammatically-oriented article.

*Archiving* is the activity of preserving primary records, which is especially important in the case of languages for which there's very little documentation.

While the best chance of accurate and complete documentation is done *in situ* in the field, involving many speakers, and multiple contexts, even an imperfect elicitation with a single speaker in a linguistics office can have extraordinary documentary value. As Himmelmann (1998) establishes in his classic article, there is an important distinction between documentation and description, which refers to analysis, whether of the phonology, the morphology, their interaction, or other aspects of the language. While theory might guide the documentation process in the sense of predetermining the questions that are asked and the issues which are investigated, it is particularly in producing the descriptive analysis that questions of theory move to center stage, thereby producing the following paradox concerning the role of theory: One needs perhaps something like Dixon's (1997:128) 'basic linguistic theory' to know what to look for and how to interpret what one finds, but one's theory should not be so specific or idiosyncratic that each linguist comes up with a different analysis of the same facts, or the same analysis of different facts. In other words, when we go to compare grammatical properties across languages, we want to be sure that we are typologizing linguistic systems, and not linguists.

Finally, archiving is the third component which has received a lot of recent attention in documentary/field linguistics. While much of the emphasis is on audio and video recording, the same archiving needs apply to handwritten transcriptions. A lot of us have boxes of fieldnotes and 3x5 cards which are yet to be scanned and made available to the linguistic community at large.

As indicated in (2), archiving is felt to be most important where documentation and hence descriptions are scarce. I have to admit that even before the current movement(s), I have always been drawn to study languages on which there has been limited research. The one major exception in my career is Luganda, which was already one of the best studied Bantu languages at the time that Francis Katamba and I decided to take a closer look. We would like to think that the numerous studies we did made additional contributions both to documenting new facts and analyzing that language, particularly with an eye open to theory. However, both in the field and at home, it is possible to go out of one's way to find speakers of languages that have hardly been studied or have not been studied at all. This includes very small languages for which a single speaker has left the country to work or study, with whom a non-canonical field situation can be in part simulated in a linguistics office.

For the rest of this paper I would like to focus on one such case, Nzadi, a minority Bantu language spoken by a community of fishermen on the Kasai River. When we began our study in a field methods course in 2008, Nzadi was not only unstudied (aside from an unpublished word list), but had not even been mentioned in the literature, e.g. at ethnologue.com. Mahou (2009) has since classified it as B.865, putting it in the same group as Dzing (B.86) and a few other small, understudied languages. A descriptive sketch grammar is being planned with Thera Crane and our consultant, Simon Nsielanga Tukumu, a native of Bundu, which is marked on the following map:

### Map 1

<http://www.maplandia.com/democratic-republic-congo/bandundu/bundo/>



Although Bantu, Nzadi has undergone extensive reduction and restructuring, hence my subtitle, ‘grammatical loss and innovation in Nzadi’. What I would like to show is that when languages undergo radical simplification, they not only lose things, but also get new ideas.

### 3. GRAMMATICAL LOSS AND INNOVATION IN NZADI

From a Bantu perspective, it is striking how short words are in Nzadi: Out of 1,000 lexical entries, 859 (or 85.9%) contain a monosyllabic stem, while 141 (or 14.1%) have a bisyllabic stem. While many of the latter are reduplications, borrowings or compounds, the monosyllabic stems often derive transparently from longer Proto-Bantu (PB) forms via loss of consonants and vowels, but with preservation of the tones, e.g. Proto-Bantu trisyllabic \*L-H-L becomes Nzadi monosyllabic LHL:

- (3) PB \**mù-ánà* > *mwǎàn* ‘child’  
 PB \**bù-átò* > *wǎàr* ‘canoe’  
 PB \**lù-bókò* > *lwǒ* ‘arm, hand’

Unlike PB, Nzadi sentences are largely isolating: verbs are generally monosyllabic, e.g. /bùl/ ‘hit’, /túm/ ‘send’, while nouns may or may not have singular and plural V- or N- prefixes: *sìŋ* ‘net(s)’, *ò-káàr* ‘woman’, *à-káàr* ‘women’, *n-tsúr* ‘animal(s)’. Although these derive from PB noun classes, as indicated in (4), the only true noun class agreement which remains (with some obscuration) is the ‘connective’ (genitive) marker /é/ used with all PB noun classes except classes 1 and 9:

(4)	PB	singular		plural	PB
	1, 3	ò-	—	è-	4
	5	ì-	—	à-	2, 6
	7	è-	—	N-	10
	9	N-			

- |     |                   |                  |                   |                    |
|-----|-------------------|------------------|-------------------|--------------------|
| (5) | <i>ò-lúm mǐ</i>   | ‘my husband’ (1) | <i>à-lúm é bí</i> | ‘our husbands’ (2) |
|     | <i>ò-sìm è mí</i> | ‘my rope’ (3)    | <i>è-sìm è mí</i> | ‘my ropes’ (4)     |
|     | <i>ì-kàŋ è mí</i> | ‘my spear’ (5)   | <i>à-kàŋ è mí</i> | ‘my spears’ (6)    |
|     | <i>è-túŋ é mí</i> | ‘my fly’ (7)     | <i>n-túŋ é mí</i> | ‘my flies’ (10)    |
|     | <i>ŋ-gòm mǐ</i>   | ‘my drum’ (9)    | <i>ŋ-gòm è mí</i> | ‘my drums’ (10)    |

Otherwise Nzadi has dropped the noun class system in favor of singular/plural and human/non-human marking on noun phrase elements. As seen in (6), only pronouns, which are independent words, show full singular/plural and [ $\pm$ human] agreement. Demonstratives distinguish [ $\pm$ human] only in the plural, while only a subset of adjectives show singular-plural agreement (others are invariant):

(6)	singular		plural	
	human	non-human	human	non-human
pronouns	<i>ndé</i>	<i>nǎ</i>	<i>bǎ</i>	<i>mǎ</i>
demonstratives (e.g. ‘this’)	<i>ná-pè</i>		<i>bá-pè</i>	<i>má-pè</i>
adjectives (some; e.g. ‘big’)	<i>ò-nân</i>		<i>à-nân</i>	

Crucially, although main clauses have the standard Bantu SVO structure, there is no subject-verb agreement: *mùùr á túl* ‘the person has arrived’, *bààr á túl* ‘the people have arrived’ (*á* = perfect marker). In addition, pronouns are independent words and have the same shape, whether used as subject, object, or possessive. As seen in (7), the human plural pronouns have fused the PB class 2 prefix *\*ba-*, while the non-human 3rd person pronouns have fused PB class 5 *\*lɿ-* (sg.) and class 6 *\*ma-* (pl.):

(7)	singular	plural
1st person	<i>mǐ</i>	<i>bǐ</i>
2nd person	<i>yǎ</i>	<i>byěñ</i>
3rd person [+human]	<i>ndé</i>	<i>bǎ</i>
3rd person [-human]	<i>nǎ</i>	<i>mǎ</i>

With this background on grammatical loss, we can now consider the most interesting grammatical innovation that we found in Nzadi: the obligatory expression of the subject in a non-subject relative clause (NSRC). In Bantu, as elsewhere, it is not uncommon to find postposition of the subject of a (non-subject) relative clause, as in (8a), a structure which is found also in nearby languages, e.g. Sakata (Monse 1987:104) and Yansi (Salikoko Mufwene, personal communication):

- (8) (a) *mwàán (nà ɲgè) ò mǎn àkáàr* ‘the child that the women saw’  
 child (that which) PAST see women  
 = (b) *mwàán (nà ɲgè) àkáár ↓ ó mǎn bǎ* ‘the child that the women saw’  
 child (that which) women PAST see they  
 (c) *\*mwàán (nà ɲgè) àkáár ↓ ó mǎn*

In (8b), however, we see that the subject noun can appear in preverbal position, as in main clauses, but only if an agreeing pronoun (*bǎ* ‘they/them’) cooccurs after the verb. The sentence in (9a) shows the same VS structure when the subject is a pronoun, which, however, as seen in (9b), cannot occur preverbally:

- (9) (a) *mwàán (nà ɲgè) ò mɔ́n bɔ̃* ‘the child that they saw’  
 (b) \**mwàán (nà ɲgè) bɔ̃ o mɔ́n bɔ̃*  
 (c) \**mwàán (nà ɲgè) bɔ̃ ò mɔ́n*

A NSRC such as in (9c) is thus ungrammatical for two reasons: the subject is not expressed postverbally and a pronoun cannot occur preverbally.

The same facts are observed when an adjunct is relativized:

- (10) (a) *èsúú (nà ɲgè) ò mɔ́n àkáár mwǎàn* ‘the day that the women  
 day (that which) PAST see women child saw the child’  
 = (b) *èsúú (nà ɲgè) àkáár ò mɔ́n bɔ̃ mwǎàn*  
 (c) \**èsúú (nà ɲgè) àkáár ò mɔ́n mwǎàn*  
 (d) *èsúú (nà ɲgè) ò mɔ́n bɔ̃ mwǎàn* ‘the day that they saw  
 the child’  
 (e) \**èsúú (nà ɲgè) bɔ̃ ò mɔ́n (bɔ̃) mwǎàn*

The ungrammaticality of (10c) again shows the obligatoriness of the VS structure, while (10e) is again ungrammatical because a pronoun cannot appear preverbally.

The questions that need to be asked, therefore, are: What is going on here? Why is it the way it is? What can we relate it to? Even though there is an agreement between the preverbal noun and postverbal pronoun in (8b) and (10b), the assumption is that the ‘real’ subject is what appears postverbally in the VS structure. This then raises the question of what the preverbal NP is doing there. It turns out, first, that the same options are available in non-subject WH- and cleft sentences, and, second, that there is no difference in meaning between (8a,b) or (10a,b): attempts to attribute a topic or focus function to the preverbal subject fail to account for the fact that both (8a, 10a) and (8b, 10b) are available with ‘even focus’ or when contrastive focus is placed on either the subject, the object, or the verb in the relative clause. The only other clue, equally puzzling, is seen in (11).

- (11) (a) *èsúú (nà ɲgè) àkáár ò mɔ́n bɔ̃ ńdé* ‘the day that the women  
 day (that which) women PAST see they him saw him’  
 (b) \**èsúú (nà ɲgè) ò mɔ́n àkáár ńdé*  
 (c) *èsúú (nà ɲgè) ò mɔ́n bɔ̃ ńdé* ‘the day they saw him’  
 (d) *àkáár ↓ ó pé mwǎàn n ɔ̃* ‘the women gave the  
 women PAST give child it child it’

When the subject is a noun and the object a pronoun, the subject noun can precede the verb, as in (11a), but not follow as in (11b). (11c) shows that there is nothing wrong when both the subject and object are pronouns, and (11d) shows that a noun + pronoun sequence is grammatical in a double object construction.

The second question is: Why does Nzadi do this? The question is not why the language allows a VS order in NSRCs, rather why the VS is obligatory even when the subject is expressed as a full noun before the verb. Synchronically, how should it be modeled? Diachronically, how did it come into existence? Since REL-V-S is the ‘type le plus représenté au relatif objectif à sujet lexical’ in Bantu

(Nsuka Nkutsi 1982:77), it makes sense to assume that (8a) is an older structure than (8b), which is also much more rare—which naturally brings us to the third question: What can we relate the obligatory, recapitulated VS pronouns to? Within Bantu striking parallels are, to my knowledge, found only in zone P Bantu languages such as Lomwe and Makua, which have the typical Bantu prefixal subject marking, but also place after the verb a pronominal subject of a NSRC with a possessive form (Nsuka Nkutsi 1982:72-73). The Lomwe NSRC in (12) seems quite parallel to Nzadi except for the more conservative Bantu agglutinative structure and the possessive form of the postverbal pronoun:

- (12) *mutchu owo mu-hi-na-mu-suwel-anyu* ‘the man that you do not  
man that you.PL-NEG-TM-him-know-your.PL know’ (TM = tense marker)

While the subject of the relative clause is marked both by the subject prefix and the possessive pronoun, the subject prefix of the NSRC instead agrees with the relativized noun in Makua. The following NSRC is from van der Wal (in press):

- (13) *ekamisá e-pasar-aly-áaka* ‘the shirt that I ironed’  
shirt it[cl.9]-iron-PERF.REL-my

Since it appears to have the structure ‘the shirt that ironed my’, van der Wal proposes that the NSRC is a ‘participial modifier’. As pointed out with respect to the forms in (7), Nzadi, unlike canonical Bantu, has a single set of forms which serve in a pronominal function. There is in fact no evidence that VS pronouns such as in (8b) and (10b) are possessives, e.g. they cannot be preceded by the connective /é/. Nsuka Nkutsi provides possible historical scenarios for the development of subject-marking VS possessive pronouns, which it should be noted often resemble independent pronouns (‘substitutifs’) in Bantu (Kamba Muzenga 2003).

The second parallel within Bantu is slightly better established, but equally intriguing. A number of Bantu languages repeat the subject pronominally in a negative clause (Nsuka Nkutsi 1982:74-76; Devos & van der Auwera 2009). This is seen in the following affirmative/negative pairs in Gweno (Philipsson 1993):

- |      |                        |   |                              |
|------|------------------------|---|------------------------------|
| (14) | affirmative            | negative  |                              |
|      | <i>ní-le-m-bón-íre</i> | <i>ní-le-m-bón-íre jí</i>                       | ‘I saw/didn’t see him’       |
|      | <i>kú-le-m-bón-íre</i> | <i>kú-le-m-bón-íre pfó</i>                      | ‘you sg. saw/didn’t see him’ |
|      | <i>á-le-m-bón-íre</i>  | <i>á-le-m-bón-íre wé</i>                        | ‘s/he saw/didn’t see him’    |
|      | <i>fú-le-m-bón-íre</i> | <i>fú-le-m-bón-íre fwé</i>                      | ‘we saw/didn’t see him’      |
|      | <i>mí-le-m-bón-íre</i> | <i>mí-le-m-bón-íre mwé</i>                      | ‘you pl. saw/didn’t see him’ |
|      | <i>Bá-le-m-bón-íre</i> | <i>Bá-le-m-bón-íre Bó</i>                       | ‘they saw/didn’t see him’    |
|      | SM-TM-him-see-TM       | SM <sub>i</sub> -TM-him-see-TM PRO <sub>i</sub> | (SM = subject marker)        |

The recapitulative postverbal subject markers have been variously identified as possessive, independent, demonstrative, or reflexive/logophoric pronouns



according to their morphology also in Grassfields Bantu languages, where such negative structures have also been reported (Asongwed 1980, Watters 2003:251, Leroy 2003:329-330, Mihas 2009). The question is whether there is any link between the (possibly possessive) pronouns which recapitulate the subject in Nzadi, Lomwe and Makua, and what has been reported to occur in negative constructions both within Narrow Bantu and Grassfields Bantu.

#### 4. CONCLUSIONS

The above summarizes what I plan to develop for my presentation at LDLT2. As I stated in section 2, part of my goal is to demonstrate the value of documenting languages under restricted circumstances, e.g. *ex situ* and, in this case, with only one available language consultant. There is no question that a full-scale study of the phenomena discussed in section 3 would yield a richer, more complete picture of what is going on. Given the obvious limitations of the work, we cannot take as definitive, for example, that (8a) and (8b) are truly synonymous and interchangeable, a judgment which we however confirmed from several angles, on different occasions, and always with the same result—or, whether careful documentation of multiple speakers in different discourse settings would reveal a pragmatic distinction between the two structures. Still, in the study it was possible to document judgments of grammaticality, variability, and, in some cases inconsistencies which suggest both certain interpretations, as well as avenues for future investigation. The detailed documentation that was possible has established for sure that there is something interesting and elusive going on, which in turn led me to look to other Bantu languages where there appear to be related, if not identical VS structures. Among the questions facing us now are: Why do Bantu languages have constructions where the subject is expressed both before and after the verb (as a pronoun, sometimes possessive)? Are there comparable phenomena to be found in languages from elsewhere? Whatever the answers, I hope to have demonstrated that good—and mysterious—things do come in small languages.

#### REFERENCES

- Asongwed, Tah. 1980. Sentence Negation in ngamambo in L'expansion bantoue. In Luc Bouquiaux (ed.), *L'expansion bantoue*, vol. 2, 555-561. Paris: SELAF.
- Devos, Maud & Johan van der Auwera. 2009. Jespersen cycles in Bantu: double and triple negation. Ms. Royal Museum for Central Africa & University of Antwerp.
- Dixon, Robert M. W. 1997. *The rise and fall of languages*. Cambridge University Press.
- Himmelman, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36.161-195.

- Hyman, Larry M. 2001. Field work as a state of mind. In Paul Newman & Martha Ratliff (eds), *Linguistic fieldwork*, 15-33. Cambridge University Press.
- Kamba Muzenga, J.G. 2003. *Substitutifs et possessifs en bantou*. Louvain: Peeters.
- Leroy, Jacqueline. 2003. *Grammaire du Mankon: Langue du Bantou des Grassfields parlée dans la province Nord-Ouest du Cameroun*. Thèse d'Etat, Université de Paris.
- Mahou, Jouni. 2009. NUGL Online. The on line version of the New Updated Guthrie List, a referential classification of the Bantu languages, June 4, 2009.  
<http://goto.glocalnet.net/mahopapers/nuglonline.pdf>
- Mihas, Elena. 2009. Negation in Metta. *Rice Working Papers in Linguistics* 1.197-222.
- Monse Ikamba. 1987. *Elements de grammaire sakata (BC34) (parler de Mongobele) (Rep. du Zaire)*. Bandundu, Republique du Zaire: CEEBA.
- Philippon, Gérard. 1993. Les négatifs en gweno (bantou, E 65). Paper presented at CALL, Leiden, August 29-31, 1993.
- Prince, Alan (2007). In pursuit of theory. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*, 33-60. Cambridge University Press.
- van der Wal, Jenneke. In press. Makhuwa non-subject relatives as participial modifiers. *Journal of African Languages & Linguistics*.
- Watters, John R. 2003. Grassfields Bantu. In Derek Nurse & Gérard Philippon (eds), *The Bantu Languages*, 225-256. London: Routledge.