Agentive Nominalizations in Gĩkũyũ and the Theory of Mixed Categories

Joan Bresnan & John Mugane

Revised and corrected May 28, 2006

Mixed categories[1] are constructions which combine the syntactic and morphological properties of two distinct categories, such as noun and verb, while being headed by a single word. These constructions challenge two basic principles of syntax—endocentricity and lexical integrity:

(1) a. **Endocentricity:** every phrasal projection has a unique lexical head which determines its categorial properties.

b. **Lexical integrity:** every lexical head is a morphologically complete word formed out of different elements and by different principles from syntactic phrases.

The paradigm examples of mixed categories have often been taken to be verbal forms such as gerunds and infinitives. One example is the English construction in (2).

(2) a politician's reportedly not telling the public the truth

[1]In order to make examples morphologically transparent, we have transcribed certain vowels separately between morpheme boundaries even though some of these vowels are elided or coalesced in speech. Our segmental transcriptions follow Mugane 1996 and 1997. The seven vowels transcribed *i, ĩ, e, a, o ũ, u* have the approximate respective values [i, e, ɛ, a, ɔ, o, u], each of which can be short or long. Long vowels are indicated by doubling. *m, n, ny, ng'* represent labial, coronal, palatal, and velar nasals, respectively. *mb, nd, nj, ng* are the corresponding prenasalized stops, and *b, th, c, g* the corresponding fricatives. The glosses use Arabic numerals to represent noun classes and small Roman numerals to represent values in the category of person. Abbreviations are NOM/nominalizer, ASSOC/associative, (adnominal) particle, DEM/demonstrative, APPLIC/applicative, RECIP/reciprocal, NEG/negation, FUT/future, OBJ/object, REL/relativizer, INT/intensifier, PERF/perfect, SUBJ/subject, REFL/reflexive, INTJ/interjection, FV/final vowel, SG/singular, PERF.PART/perfect participle, HAB.PART/habitual participle.

In (2) the genitive NP, otherwise exclusively a constituent of noun phrases in English, cooccurs with constituents otherwise exclusively found in verb phrases—the double NP complements.[2] The standard sentential negation *not* and the preverbal adverb in this example are also typical VP constituents. Thus the construction mixes together properties found exclusively in verb-headed structures like (3) with those found in purely nominal structures like (4):

(3) He is reportedly not telling the public the truth.

(4) a politician's reported (non-)telling(s) of the truth to the public[3]

Another example is the Italian *infinito sostantivato* (what Zucchi 1993 calls "VP-infinitival NPs"). In (5)–(6) determiners, possessives, and qualifying adjectives—NP/DP constituents—appear before the infinitive, while direct object NPs and adverbs—VP constituents—appear after the infinitive (Zucchi 1993):

(5) *il suo continuo momorare parole dolci*
the his/her continual whisper.INF words sweet
'his continual whispering of soft words' (Zucchi 1993: 239)

(6) *il suo scribere quella lettera improvvisamente*
the his/her write.INF that letter suddenly
'his suddenly writing that letter' (Zucchi 1993: 54)

The purely nominal infinitives (what Zucchi 1993 calls "N-infinitival NPs") take adjectives and not adverbs as modifiers; compare the nominalization in (7):

(7) *la cessazione improvvisa/*improvvisamente delle ostilitá*
the cessation sudden/suddenly of the hostilities
'the sudden cessation of the hostilities' (Zucchi 1993: 223)

[2]Double NP complements also occur in ellipsis constructions having VP antecedents, such as the sequence of NPs [*their cats*] [*cabbage*] following *than* in *The survey revealed that more of the children gave their dogs spaghetti than their cats cabbage.*

[3]The nominal gerund differs morphologically from both the verbal gerund and the participle in allowing negative prefixation by *non-* and plural suffixation, which are impossible with participles and verbal gerunds. In addition its syntactic complement type (the *of* phrase) is also characteristic of relational nouns underived from verbs; the latter take PPs, like the *of* phrase in *Her picture of Mary, *Her picture Mary.*

The *infinito sostantivato* shows its mixed properties in being able to take both adjectives and adverbs at the same time:

(8) *il suo continuo eseguire la canzone impeccabilimente*
the his/her continual perform.INF the song impeccably
(Zucchi 1993: 55)
'his continually performing the song impeccably'

However, the mixed verbal constructions that we see in English and Italian have a special property which is not true of mixed category constructions in Gĩkũyũ: their meanings belong to the same semantic type as those of clauses headed by verbs. Zucchi (1993: 251ff) argues that the mixed Italian constructions (such as (5), (6), and (8)) denote proposition-like entities, while N-infinitival NPs like (7) denote events. He argues that a similar difference appears with English verbal and nominal gerundive constructions (*his performing the song* vs. *his performing of the song*) (Zucchi 1993: 67–71).

Because the meanings of infinitives and gerunds are of the same semantic types as those of verbs, these kinds of mixed categories can be regarded simply as inflectional subtypes of the base lexical category. The problem of mixed categories then appears to be primarily a syntactic one of correlating the morphology of the head with the categorially mixed syntax. However, Gĩkũyũ shows us that mixed categories can be derived not only by inflectional morphology, but by morphology which fundamentally changes the category of lexical meaning type—that is, by what is considered to be classically lexical derivational morphology. Gĩkũyũ has deverbal agentive nominalizations analogous to the English example in (9), except that they are fully grammatical. An example is given in (10).[4]

(9) *the driver a rusty truck to Arizona reluctantly

(10) *ũyũ mũ-thĩĩnj-i mbũri ũũru*
1.DEM 1-slaughter-NOM 10.goat badly
'this bad goat slaughterer'; lit.: 'this slaughterer goats badly'

The important properties of these Gĩkũyũ constructions have been developed in Mugane (1996, 2003).[5] In what follows we review both the morphosyn-

[4]Example (9) is based on Ackema and Neeleman (2001), and its implications are discussed further in Sections 4 and 5 below.

[5]Mugane 2003 also provides tonal transcriptions, which unfortunately could not be included in the present study.

tax of such agentive nominalizations in Gĩkũyũ and the range of available analyses, in order to develop an explanation within the LFG framework which adheres to principles (1a,b).

1 Gĩkũyũ agentive nominalizations

Although prototypical referents of these nominalizations are agents (e.g. *mũ-in-i*, 1-sing-NOM, 'singer'), they may also have other roles, such as instrument *gĩ-thĩĩnj-i*, 7-slaughter-NOM, 'something to slaughter with', *i-thĩĩnj-i*, 8-slaughter-NOM, 'things to slaughter with' (plural).[6] We use the term 'agentive' nominalization with the understanding that agents are only the typical and not the exclusive referents of these nominals.

1.1 Morphology

Gĩkũyũ agentive nominalizations are illustrated in (11):

(11) a. *mũ-in-i*
1-sing-NOM
'singer'

b. *mũ-thĩĩnj-i*
1-slaughter-NOM
'slaughterer'

c. *a-ndĩk-i*
2-write-NOM
'writers'

Note that these nominalizations bear noun class markers, which are the prefixes glossed by Arabic numerals in (11). These clearly mark (11a-c) as belonging to the inflectional class of nouns in Bantu. Other categories have concordial class marking prefixes, but they differ in shape in some classes (Mugane 1997: 26–27). For example, the class 8 prefix *ci-,i-* does not appear on adjectives and adnominal verbs, which instead mark class 8 by prenasalizing the initial consonant of the base. The subject prefixes of finite verbs differ from nouns in

[6]To account for this fact in English, Rappaport Hovav and Levin 1992 use the term '*-er* nominal', but it obviously lacks crosslinguistic identifiability.

classes 1, 3, 4, 5, 9, and 10. In (12) the same stem is shown with four different noun class markers:

(12) a. *mũ-thuur-i*
1-select-NOM
'selector (human)'

b. *gî-thuur-i*
7-select-NOM
'selector (augmentative/derogatory)'

c. *ma-thuur-i*
6-select-NOM
'selectors (collective)'

d. *tũ-thuur-i*
13-select-NOM
'selectors (diminutive/ameliorative)'

The noun class markers indicate the class with which all nominal modifiers and predicates must agree. For example, if a noun phrase headed by an 'augmentative/derogatory' class noun (12b) appears as the subject of a verb, the verb must show class 7 agreement with its subject marking prefix; a quantifier phrase modifying the subject nominal must also show class 7 agreement.

The noun class marker is prefixed to a verb stem to which a nominalizing suffix has been attached. The agentive suffix *-i* in these nominalizations is one of a series of suffixal nominalizers (Mugane 1997: 88–89). Several are illustrated below with the verb stem for 'slaughter'. (Many other deverbal nominalizing suffixes occur; for example, distinct suffixes exist for deverbal occasions, abstract concepts, and states.)

(13) Nominalizations of the verb *thĩĩnja* 'slaughter':

Nominalization	Gloss	Type
mũ-thĩĩnj-i	'slaughterer' (class 1)	agentive
mũ-thĩĩnj-ĩre	'manner of slaughter' (class 3)	manner
gĩ-thĩĩnj-ĩro	'slaughter location' (class 7)	location

The forms of the nominalizing suffixes are generally not possible as verb desinences in Gĩkũyũ, further evidence that the derived forms do not belong to the

inflectional class of verbs.[7] The agentive suffix requires that the verb stem have an agentive role semantically, accounting for the fact that nonagentive verbs 'be' and 'have' are semantically incompatible with agentive nominalization:

(14) a. * *mũ-korw-i*
1-be-NOM
'a "be-er," one who is'

b. * *mũ-rĩ-i*
1-have-NOM
'a "hav-er," one who has'

The verbal base of the nominalization may undergo various stem derivation processes exclusive to verbs, including the reduplication, applicativization, and reciprocalization illustrated in (15a–c):[8]

(15) a. *mũ-rũgarũg-i*
1-jump.jump-NOM
'one who jumps repeatedly halfheartedly'

[7]However, *-i* can be used to turn a verb into a habitual participle which is used with the adjective class prefixes as an adnominal modifier, as discussed below in section 2.

[8]The verbal base of the nominalization cannot be inflected for negation, tense, or aspect (although unnominalized verbs may be):

(i) * *mũ-ti-on-i*
1-NEG-see-NOM
'one who does not see'

(ii) * *mũ-ka-on-i*
1-FUT-see-NOM
'one who will see (habitual)'

This restriction cannot be attributed to a ban on prefixal verbal morphology on the verbal base, however, because an aspectual suffix of verb stems is also excluded—

(iii) * *mũ-on-ag-i*
1-see-HAB-NOM
'one who sees (habitual)'

—and because a reflexive prefix may appear:

(iv) *mũ-ĩ-on-i*
1-10.REFL-see-NOM
'one who sees himself/herself (a braggart)'

b. *a-ndĩk-ir-i*
2-write-APPLIC-NOM
'those who write for/to (others)'

c. *a-ndĩk-án-i*
2-write-RECIP-NOM
'those who write each other'

In sum, these agentive nominalizations consist of a verbal base which is nominalized by an agentive suffix and prefixed by a noun class marker. The base undergoes a subset of verbal morphological processes, including verbal extension by suffixation, reduplication, and reflexive prefixing. The meaning, inflectional class, lexical category, and morphological type of the nominalization tell us that it is a deverbal noun.

1.2 NP constructions

Agentive nominalizations may head purely nominal syntactic phrases, as in (16):

(16) a. *mũ-in-i* *w-a* *i-tũũra*
1-sing-NOM 1-ASSOC 5-settlement
'singer of the settlement'

b. *mũ-in-i* *w-a* *nyĩmbo*
1-sing-NOM 1-ASSOC 10.song
'singer of songs'

While the verb 'sing' takes a direct NP object, the agentive nominalization 'singer' in (16b) takes an associative phrase expressing the semantic role of the object. An associative phrase is an adnominal phrase headed by a particle *-a* 'of' which bears a concordial prefix agreeing in noun class with the nominal it modifies. The same associative marker marks nominal complements and nominal adjuncts, as illustrated in (16a,b). Complement associative phrases (16b) appear in exactly the same positions as the adjunct phrases (16a) relative to other adnominal constituents (Mugane 1996). For example, in the unmarked order of NP-constituents shown abstractly in (19), both are separated from the head by the demonstrative:

(17) a. *mũ-in-i* *ũyũ* *w-a* *i-tũũra*
1-sing-NOM 1.DEM 1-ASSOC 5-settlement
'this singer of the settlement'

b. *mũ-in-i* *ũyũ* *w-a* *nyĩmbo*
1-sing-NOM 1.DEM 1-ASSOC 10.song
'this singer of songs'

Hence the associative phrase interpreted as a complement to the head nominal is probably an argument adjunct—an optional adjunct that is interpreted with respect to a specific argument role of the head nominal, much as the passive agentive phrase has been analyzed (e.g. by Alsina 1996).

Demonstratives, possessive pronouns, adjectives, and relative clauses may also modify an NP headed by an agentive nominalization. The head N is NP-initial, preceding all of these constituents except for the determiner, which may optionally appear in initial position when focused:

(18) a. *mũ-in-i* *ũyũ* , *ũyũ* *mũ-in-i*
1-sing-NOM 1.DEM 1.DEM 1-sing-NOM
'this singer'

b. *mũ-in-i* *w-itũ*
1-sing-NOM 1-our
'our singer'

c. *a-in-i* *a-nene*
2-sing-NOM 2-big
'big singers'

d. *a-in-i* *a-rĩa* *ũ-ĩ*
2-sing-NOM 2-REL 2.SG.SUBJ-know
'the singers whom you know'

The pragmatically unmarked order of nominal dependents is shown in (19) (Mugane 1996: 88).[9]

(19) `N < Dem < PossPron < QP < AP < AssocP`

[9]All of these nominal constituents, including the head N, are optional. Omission of the head results in a null anaphoric interpretation.

Other word orders are possible, but are marked with pauses. These word order generalizations are exactly the same in NP constructions headed by underived nouns. Compare the unmarked orders in (20a,b):

(20) a. [*mũ-end-i*]$_{\text{N}}$ *uyu* *w-a* *a-ndũ*
1-love-NOM 1.DEM 1-ASSOC 2-person
'this lover of people'

b. [*nyũngũ*]$_{\text{N}}$ *ĩno* *y-a* *u-cũrũ*
9.pot 9.DEM 9-ASSOC 14-porridge
'this pot of porridge'

The same nominal modifiers may occur when the nominalized verb bears a reflexive prefix or a verbal extension, as exemplified by (21) and (22), respectively:[10]

(21) *mũ-ĩ-rut-i* *ũyũ*
1-10.REFL-see-NOM 1.DEM
'this one who sees himself/herself'

(22) *mũ-hũr-an-i* *ũyũ*
1-fight-RECIP-NOM 1.DEM
'this one who fights with others'

Just as the internal structure of these agentive phrases is typical of NPs, so is their external distribution. They may be subjects or objects of verbs or prepositional objects, they may induce noun class concord with their matrix verbs, and they may be clefted and relativized—all properties of NPs, but not of nonnominal categories such as VPs or CPs in Gĩkũyũ. (See Mugane 1996 for detailed exemplification.)

[10]As noted by Mugane (1996: 104–5) a pronominal object marker can be prefixed to the verb stem, but cannot cooccur with nominal modifiers. This could be explained if pronominal object prefixation turns out to be a property of the habitual participle, an adnominal form of the verb which resembles the agentive nominalization (Section 2). However, we have not yet been able to find clear evidence for or against this hypothesis because the class 8 concordial morphological distinction between nouns and participles appears to be neutralized when the participle bears an object marker, as discussed below.

1.3 Mixed NP/VP constructions

While the preceding examples of agentive nominalization constructions reveal an NP headed by a deverbal noun, the same agentive nominalizations also appear in mixed category constructions, as shown in (23):

(23) a. [*mũ-thĩĩnj-i*]$_{\text{N}}$ [*mbũri*]$_{\text{NP}}$ [*wega*]$_{\text{ADV}}$ *w-a* *Nairobi*
1-slaughter-NOM 10.goat 1.well 1-ASSOC N.
'a good goat slaughterer from Nairobi'
Lit.:'(a) slaughterer goats well from Nairobi'

b. [*mũ-in-ĩr-i*]$_{\text{N}}$ [*a-ndũ*]$_{\text{NP}}$ [*nyĩmbo*]$_{\text{NP}}$ *ũyũ*
1-sing-APPLIC-NOM 2-person 10.song 1.DEM
'this singer of songs for people'
Lit.: 'this singer people songs'

c. [*mũ-in-i*]$_{\text{N}}$ [*wega*]$_{\text{ADV}}$ *ũ-rĩa* *mũ-nene*
1-sing-NOM well 1-REL 1-big
'the one who sings well who is big'
Lit.: '(the) singer well who is big'

The Gĩkũyũ constructions in (23a-c) consist of the head, which is an agentive nominalization, immediately followed by a sequence of verbal dependents—a direct object and adverb in (23a), two NP objects in (23b), and an adverb in (23c)—followed in turn by nominal dependents—the associative ('of' phrase) adnominal modifier in (23a), the demonstrative in (23b), and a relative clause in (23c). Elsewhere in Gĩkũyũ, double NP complements occur exclusively in VPs and certain adverbial adjuncts are not found as the immediate constituents of NPs or DPs.

The semantic types of adverbial modifiers include manner ('skillfully', 'cleverly', 'quickly', 'slowly', 'carefully'), duration ('for a long time'), temporal ('early'), evaluation ('badly', 'well'), and intensity ('very', 'totally').[11] They can be expressed by PPs ('with knowledge'), verbal phrases ('caring for them'), a small closed class of adverbs, and interjective particles. Examples of an intensifier and emphatic interjection are given in (24):

[11]Thus they are not semantically limited to the types that have fallen in the domain of verbal case assignment in case-marking languages (cf. Wechsler and Lee 1996, Przepiórkowski 1999, Hanjung Lee 1999a,b).

(24) a. *mũ-kir-i* *mũno*
1-quiet-NOM very
'one who is very silent'

b. *mũ-kir-i* *ki*
1-quiet-NOM INTJ
'one who is totally silent'

Not only do constituents of the VP occur in these mixed categories, they must occur in exactly the same order as in sentence VPs. The order of VP constituents in a simple sentence is shown in (25) (Mugane 1996: 142):

(25) `Verb < Indirect Object < Direct Object < Target Locative < Manner Adverbial < Setting Locative < Temporal Adverb`

For example, an adverb must follow any NP objects in the mixed category construction:

(26) a. [*mũ-in-ĩr-i*]$_{\text{N}}$ [*a-ndũ*]$_{\text{NP}}$ [*nyĩmbo*]$_{\text{NP}}$ [*wega*]$_{\text{ADV}}$
1-sing-APPLIC-NOM 2-person 10.song well
'one who sings songs for people well'

b. *[*mũ-in-ĩr-i*]$_{\text{N}}$ [*a-ndũ*]$_{\text{NP}}$ [*wega*]$_{\text{ADV}}$ [*nyĩmbo*]$_{\text{NP}}$
1-sing-APPLIC-NOM 2-person well 10.song

c. *[*mũ-in-ĩr-i*]$_{\text{N}}$ [*wega*]$_{\text{ADV}}$ [*a-ndũ*]$_{\text{NP}}$ [*nyĩmbo*]$_{\text{NP}}$
1-sing-APPLIC-NOM well 2-person 10.song

The same is true in the corresponding sentence VPs:

(27) a. *nĩ-a-a-in-ĩr-a* [*a-ndũ*]$_{\text{NP}}$ [*nyĩmbo*]$_{\text{NP}}$ [*wega*]$_{\text{ADV}}$
FOC-iii.SG.SUBJ-PERF-sing-APPLIC-FV 2-person 10.song well
'she/he has sung songs for people well'

b. **nĩ-a-a-in-ĩr-a* [*a-ndũ*]$_{\text{NP}}$ [*wega*]$_{\text{ADV}}$ [*nyĩmbo*]$_{\text{NP}}$
FOC-iii.SG.SUBJ-PERF-sing-APPLIC-FV 2-person well 10.song

c. **nĩ-a-a-in-ĩr-a* [*wega*]$_{\text{ADV}}$ [*a-ndũ*]$_{\text{NP}}$ [*nyĩmbo*]$_{\text{NP}}$
FOC-iii.SG.SUBJ-PERF-sing-APPLIC-FV well 2-person 10.song

Similarly, the applied beneficiary object must precede the theme object in the mixed category construction (28a), and the same holds in the VP of a sentence (28b):

(28) a. *[*mũ-in-ĩr-i*]$_{\text{N}}$ [*nyĩmbo*]$_{\text{NP}}$ [*a-ndũ*]$_{\text{NP}}$ [*wega*]$_{\text{ADV}}$
1-sing-APPLIC-NOM 10.song 2-person well
'one who sings songs for people well'
Lit.: 'one who sings songs people well'

b. *[*nĩ-a-a-in-ĩr-a*]$_{\text{V}}$ [*nyĩmbo*]$_{\text{NP}}$ [*a-ndũ*]$_{\text{NP}}$ [*wega*]$_{\text{ADV}}$
FOC-iii.SG.S-PERF-sing-APPLIC-FV 10.song 2-person well
'one who sings songs for people well'
Lit.: 'she/he has sung songs people well'

Moreover, an object NP within the agentive nominalization phrase is in complementary distribution with the reflexive prefix (29), just as it is in sentence VPs (30):

(29) a. *mũ-rut-i ci-ana Gĩ-thweri*
1-teach-NOM 8-child 7-Swahili
'one who teaches children Swahili'

b. *mũ-ĩ-rut-i Gĩ-thweri*
1-REFL-teach-NOM 7-Swahili
'one who teaches himself/herself Swahili'

c. **mũ-ĩ-rut-i ci-ana Gĩ-thweri*
1-REFL-teach-NOM 8-child 7-Swahili
Lit.: 'one who teaches himself/herself children Swahili'

(30) a. *nĩ-a-a-rut-a ci-ana Gĩ-thweri*
FOC-iii.SG.SUBJ-PERF-teach-FV 8-child 7-Swahili
'She/he has taught children Swahili'

b. *nĩ-a-a-ĩ-rut-a Gĩ-thweri*
FOC-iii.SG.SUBJ-PERF-REFL-teach-FV 7-Swahili
'She/he has taught herself/himself Swahili'

c. **nĩ-a-a-ĩ-rut-a ci-ana Gĩ-thweri*
FOC-iii.SG.SUBJ-PERF-REFL-teach-FV 8-child 7-Swahili
Lit.: 'She/he has taught herself/himself children Swahili.'

In general, then, all and only the post-head immediate constituents of VPs are possible post-head constituents of the mixed agentive nominalization phrase, and all and only the possible orderings of these VP constituents are possible orderings of the same constituents in the mixed agentive phrase.

Let us now turn from the VP-like portion of the structure to the NP-like portion. Note first that the full set of nominal modifiers is possible in the presence of the VP-style constituents:

(31) a. *mũ-thĩĩnj-i* *mbũri* *ũyũ*
1-slaughter-NOM 10.goat 1.DEM
'this goat slaughterer'

b. *mũ-thĩĩnj-i* *mbũri* *w-itũ*
1-slaughter-NOM 10.goat 1-our
'our goat slaughterer'

c. *a-thĩĩnj-i* *mbũri* *othe*
2-slaughter-NOM 10.goat 2.all
'all goat slaughterers'

d. *a-thĩĩnj-i* *mbũri* *a-nene*
2-slaughter-NOM 10.goat 2-big
'big goat slaughterers'

e. *mũ-thĩĩnj-i* *mbũri* *w-a* *gĩ-cagi*
1-slaughter-NOM 10.goat 1-ASSOC 7-village
'goat slaughterer of the village'

These nominal elements occur in the normal unmarked order (19) as well as the marked orders. For example, it is unmarked for a quantifier (QP) to precede an adjective phrase (AP) in a pure NP construction, and the same is true in the mixed NP construction:

(32) a. *a-thĩĩnj-i* *mbũri* *othe* *a-nene*
2-slaughter-NOM 10.goat 2.all 2-big
'all big goat slaughterers' (unmarked)

b. *a-thĩĩnj-i* *mbũri* *a-nene, othe*
2-slaughter-NOM 10.goat 2-big 2.all
'all big goat slaughterers' (marked)

Likewise, the demonstrative precedes other nominal modifiers in the unmarked order.[12]

A further point of interest is that all complements selected by the head must be of the same type: either verbal or nominal. The split complements in (33a) illustrate this; the beneficiary argument is a verbal complement type (applied object NP), while the patient argument is a nominal complement type (associative phrase). Because applied NPs cannot be expressed by associative phrases (Mugane 1997: 106), (33b) is also bad. Consequently, ditransitive nominalization is only possible with verbal-type (direct NP) complements, as in (33c):

(33) a. * *mũ-thĩĩnj-ĩr-i* *a-ndũ* *w-a* *mbũri*
1-slaughter-APPLIC-NOM 2-person 1-ASSOC 10.goat
'one who slaughters goats for people'

b. * *mũ-thĩĩnj-ĩr-i* *w-a* *a-ndũ* *w-a* *mbũri*
1-slaughter-APPLIC-NOM 1-ASSOC 2-person 1-ASSOC 10.goat
'one who slaughters goats for people'

c. *mũ-thĩĩnj-ĩr-i* *a-ndũ* *mbũri*
1-slaughter-APPLIC-NOM 2-person 10.goat
'one who slaughters goats for people'

This homogeneity of selected complement types is a kind of 'lexical coherence' (Malouf 1998, 2000).

Gĩkũyũ mixed categories manifest not only lexical coherence (selecting complements of uniform type) but also phrasal coherence: the verbal constituents, regardless of whether they are lexically selected by the head, cohere with each other as a constituent. Thus, while in an NP it is possible to reorder all of the

[12]One restriction, however, is that it is unacceptable to have the determiner follow an adverb (i); the NP-initial order of the determiner is preferred in this case (ii):

(i) ?? *mũ-thĩĩnj-i* *mbũri* *ũũru* *ũyũ*
1-slaughter-NOM 10.goat badly 1.DEM
'this bad goat slaughterer'

(ii) *ũyũ* *mũ-thĩĩnj-i* *mbũri* *ũũru*
1.DEM 1-slaughter-NOM 10.goat badly
'this bad goat slaughterer'

nominal dependents to produce marked orders, in the mixed construction all of the VP-type constituents must precede all of the NP-type constituents. Any ordering that interleaves the two types of constituents is disallowed. In (34a-c) we see that an associative phrase cannot interrupt a sequence of an object NP followed by Adverb:

(34) a. [*mũ-thĩĩnj-i*]$_{\text{N}}$ [*mbũri*]$_{\text{NP}}$ [*wega*]$_{\text{ADV}}$ [*w-a* *Nairobi*]
1-slaughter-NOM 10.goat 1.well 1-ASSOC N.
'a good goat slaughterer from Nairobi'

b. *[*mũ-thĩĩnj-i*]$_{\text{N}}$ [*mbũri*]$_{\text{NP}}$ [*w-a* *Nairobi*] [*wega*]$_{\text{ADV}}$
1-slaughter-NOM 10.goat 1-ASSOC N. 1.well

c. *[*mũ-thĩĩnj-i*]$_{\text{N}}$ [*w-a* *Nairobi*] [*mbũri*]$_{\text{NP}}$ [*wega*]$_{\text{ADV}}$
1-slaughter-NOM 1-ASSOC N. 10.goat 1.well

In (35a-c) a demonstrative cannot interrupt an object NP Adverb sequence:

(35) a. [*mũ-end-i*]$_{\text{N}}$ [*a-ndũ*]$_{\text{NP}}$ [*mũno*]$_{\text{ADV}}$ *ũyũ*
1-love-NOM 2-person a lot 1.DEM
'this one who loves people a lot'

b. *[*mũ-end-i*]$_{\text{N}}$ [*a-ndũ*]$_{\text{NP}}$ *ũyũ* [*mũno*]$_{\text{ADV}}$
1-love-NOM 2-person 1.DEM a lot

c. *[*mũ-end-i*]$_{\text{N}}$ *ũyũ* [*a-ndũ*]$_{\text{NP}}$ [*mũno*]$_{\text{ADV}}$
1-love-NOM 1.DEM 2-person a lot

(36a-b) shows that a relative clause cannot precede an adverb:

(36) a. [*mũ-in-i*]$_{\text{N}}$ [*wega*]$_{\text{ADV}}$ [*ũ-rĩa* *mũ-nene*]
1-sing-NOM well 1-REL 1-big
'the one who sings well who is big'
Lit.: 'singer well who is big'

b. *[*mũ-in-i*]$_{\text{N}}$ [*ũ-rĩa* *mũ-nene*] [*wega*]$_{\text{ADV}}$
1-sing-NOM 1-REL 1-big well

The impossibility of interleaving the two types of constituents holds in a specific region of structure immediately following the head N. Thus, while the demonstrative can precede the head in both the pure NP and the mixed NP/VP constructions—

(37) a. *ũyũ* [*mũ-end-i*]$_{\text{N}}$ *w-a* *a-ndũ*
1.DEM 1-love-NOM 1-ASSOC 2-person
'this lover of people'

b. *ũyũ* [*mũ-end-i*]$_{\text{N}}$ [*a-ndũ*]$_{\text{NP}}$
1.DEM 1-love-NOM 2-person
'this lover of people'

—following the head a choice must be made: either complements and modifiers will be uniformly nominal (in the permitted orders of NP-type constituents (19)), or they will be uniformly verbal (in the permitted orders of VP-type constituents (25)) until the verbal sequence is exhausted and the nominal sequence begins. In (38) the post-head demonstrative *ũyũ* marks the nominal choice-point:

(38) a. [*mũ-end-i*]$_{\text{N}}$ *ũyũ* *w-a* *a-ndũ*
1-love-NOM 1.DEM 1-ASSOC 2-person
'this lover of people'

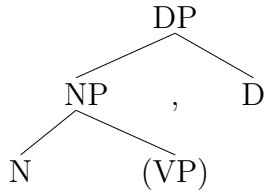
b. *[*mũ-end-i*]$_{\text{N}}$ *ũyũ* [*a-ndũ*]$_{\text{NP}}$
1-love-NOM 1.DEM 2-person
'this lover of people'

c. [*mũ-end-i*]$_{\text{N}}$ [*a-ndũ*]$_{\text{NP}}$ *ũyũ*
1-love-NOM 2-person 1.DEM
'this lover of people'

In general, then, the VP-like constituents and the NP-like constituents—regardless of whether they are selected or unselected by the head—belong to two separate, coherent regions of the structure, each subject to its own ordering constraints.

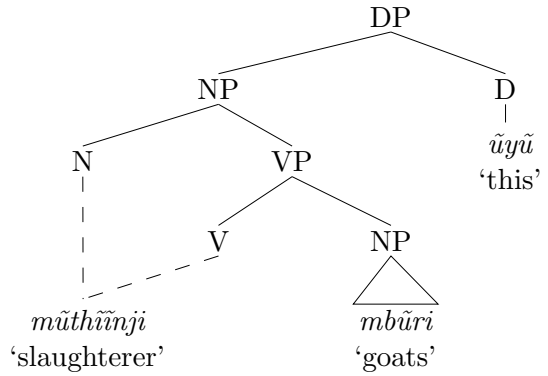
These generalizations can be explained by adding to the lexical coherence of selection for verbal or nominal complement types a requirement of phrasal coherence: the VP-style constituents within the mixed category cluster together as a unit, preventing higher nominal elements from interrupting them (Mugane 1996, Bresnan 1997):

(39)



If the VP complement in (39) is omitted, the resulting structure is that of the pure NP construction.[13] This structure can also explain without any further assumptions why the agentive nominalization in mixed categories must precede all of the other complements and modifiers (except for the optional preposing of a focused demonstrative): it occupies the typical head-initial position of Nouns in their nominal projections.[14] It remains to be answered how the two projections can share the same head, as depicted informally in (40):

(40) Gĩkũyũ mixed category:



We address this question in Section 5.

We see, then, that the internal syntax of agentive phrases seems to be grafted together from two different categorial projections sharing a single head, thereby displaying a combination of the properties displayed by VPs and NPs. Mugane 1996 shows that they also have the external syntax of nominal phrases (NPs/DPs), inducing subject or object agreement with a matrix verb and allowing extraction by clefting and relativization. These are properties not shared by VPs and CPs.

2 Alternative analyses

Two interesting alternative analyses of the mixed category facts suggest themselves. By reinterpreting the data of mixed categories as either not truly phrasal

[13]Following Mugane 1996, we assume that all of the concordial NP modifiers following the demonstrative are adjoined to DP.

[14]It is not possible to conjoin two clusters of VP-style constituents under the same head; this may be because it is not possible to conjoin two VPs in general in Gĩkũyũ.

or not truly mixed, they remove the challenge of Gĩkũyũ action nominalizations to the single projection theory. The counteranalyses also bring further properties of Gĩkũyũ and Bantu into the picture.

2.1 Synthetic compounds?

The first analysis is based on the fact that compound words in Bantu are head-initial (Mchombo 1978, Myers 1987). This fact suggests an analysis of these mixed constructions as synthetic compounds (dominated by a lexical rather than a phrasal category). After all, agentive nominalizations in English take phrasal PPs rather than direct NPs (41a), but in synthetic compounds they take a bare nominal complement unmediated by a preposition (41b):

(41) a. an eater of pumpkins, *an eater pumpkins

b. a pumpkin eater

In English the compounds can be easily distinguished from the phrases: complements appear before the head in compounds (41b), and after the head in phrases (41a). In Bantu, in contrast, complements follow the head in both compounds and phrases; one might be easily mistaken for the other.

Mugane (1996: ch. 5) argues in detail that mixed category constructions are not synthetic compounds. The complements of the agentive nominalizations may be freely modified, allowing both pre- and post-head determiners (Mugane 1996: 137–138) as in (42), and relative clauses (Mugane 1996: 154) as in (43):

(42) *mũ-end-i* [*aya* *a-ndũ*] , *mũ-end-i* [*a-ndũ* *aya*]
1-love-NOM 2.DEM 2-person 1-love-NOM 2-person 2.DEM
'a lover of these people'

(43) *mũ-thĩĩnj-i* [*ĩno* *mbũri* *njeke*]
1-slaughter-NOM 9.DEM 9.goat 9.thin
'a slaughterer of this thin goat'

The complements may also be freely coordinated (Mugane 1996: 146):

(44) *mũ-thĩĩnj-i* [*mbũri* *na* [*ngũkũ* *ici*]]
1-slaughter-NOM 10.goat and 10.chicken 10.DEM
'a slaughterer of goats and these chickens'

A pure indexical such as a deictic pronoun may be the complement (Mugane 1996: 155):

(45) *mũ-thĩĩnj-i* *ĩyo*
1-slaughter-NOM 9.DEM
'a slaughterer of that (class 9)'

These properties are not shared by lexical compounds in general (Bresnan and Mchombo 1995).[15] Thus, the mixed categories headed by agentive nominalizations in Gĩkũyũ are not synthetic compounds but phrasal constructions of syntax.

2.2 Adnominal participles?

A second alternative analysis of the mixed category facts is based on the existence of the class of adnominal participles in Gĩkũyũ. These are verbal forms which are used as modifiers of nominal heads. The participle is formed when the verb base (absent subject agreement and tense/aspect prefixes) is given a suffix (*-u* for perfect, *-e* for passive, and *-i* for habitual) and prefixed by the adjectival class markers to form an adnominal modifier. Examples are given in (46):

(46) a. (*a-ndu*) *a-thĩĩnj-u* *mbũri*
2-person 2-slaughter-PERF.PART 10.goat
'people who have slaughtered goats'

b. (*a-ndu*) *a-thĩĩnj-e*
2-person 2-slaughter-PASS.PART
'people who are operated on (by a surgical procedure)'

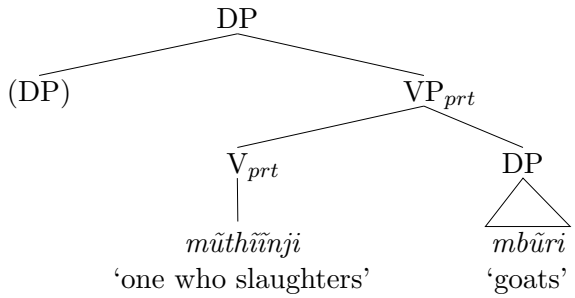
c. (*mbũri*) *thĩĩnj-e* *ni* *a-ndũ*
10.goat 10.slaughter-PASS.PART by 2-person
'goats that are slaughtered by people'

[15] All of the above properties contrast with those of a compound in Gĩkũyũ called the '[*mũ-...-a*] compound' by Mugane (1996: ch. 5). However, the Gĩkũyũ [*mũ-...-a*] forms can be loosely compounded with a single uncoordinated syntactic NP which does not begin with a determiner. This is in contrast to the much stricter constraints on phrasal recursivity shown by cognate compounds in Chichewa (Mchombo 1978, Bresnan and Mchombo 1995). See Mugane 1996 for details.

d. (*a-ndu*) *a-thĩĩnj-i* *mbũri*
2-person 2-slaughter-HAB.PART 10.goat
'people who slaughter goats'

The habitual form in (46d) resembles the agentive nominalization. Moreover, as happens with all types of NPs, the head nouns of these constructions can be dropped, yielding a definite or indefinite null anaphoric interpretation (see fn. 9). In other words, the modifying phrase can be used by itself (or in conjunction with other modifiers) as a referential expression. The structure of these examples is shown in (47):[16]

(47) Adnominal participle construction:



Could these adnominal participles be the solution to our problematic mixed category constructions? The answer is no. First, the participles bear the adjectival series of prefixes. These are formally identical to the noun prefix series except in class 8, where nouns have a prefix *ci-, i-* (depending on whether the stem begins with a vowel or not) and adjectives/participles have prenasalization of a verb-stem initial unaspirated consonant as in classes 9/10. This difference is illustrated in (48):

(48) a. *ci-ana* *ndoot-i* (< *N-rot-i*)
8-child 8.dream-PERF.PART
'children who dream (habitually)'

[16]The head position is shown in parentheses in (47). We assume that the null pronominal is not represented by a empty phrase structure category, but is functionally incorporated into the verbal morphology, where its f-structure value preempts the expression of a phrasal head in c-structure. See Bresnan and Mchombo 1987, Andrews 1990, Mugane 1996, Austin and Bresnan 1996, Bresnan 2001, and the works cited therein.

b. *ndoot-i*
8.dream-PERF.PART
'ones who dream (habitually)'

c. *i-rot-i*
8-dream-NOM
'dreamers (class 8)'

Only the adjectival prefixes may be used with adnominal verbal modifiers. Now class 8 mixed category constructions exist, headed by nominals bearing the noun prefix for class 8:

(49) *i-mũrĩk-ĩr-i* *a-ndũ* *njira* wega
8-shine-APPLIC-NOM 2-person 9.path well
'ones that illuminate paths for people well'

This fact clearly indicates that our mixed category agentive nominals are not simply adnominal participles in headless (null anaphora) constructions.

Secondly, the word order of the mixed category nominalizations differs from that of adnominal participles. Within the DP, adnominal participles occupy the same word order position as adjectives, designated 'AP' in the unmarked word order (19), repeated here:

(50) `N < Dem < PossPron < QP < AP < AssocP`

As such, they follow (in their unmarked order) all of the other types of adnominal modifiers except for associative phrases:

(51) *mũ-ndũ* *w-a-kwa* *ũ-mwe* *mũ-thĩĩnj-i* *mbũri*
1-person 1-my 1-one 1-slaughter-HAB.PART 10.goat
'my one person who is a slaughterer of goats'

They may precede or follow other APs, such as the adjective in (52):

(52) a. *mũ-ndũ* *mw-ega* *mũ-thĩĩnj-i* *mbũri*
1-person 1-good 1-slaughter-HAB.PART 10.goat
'a good person who slaughters goats'

b. *mũ-ndũ* *mũ-thĩĩnj-i* *mbũri* *mw-ega*
1-person 1-slaughter-HAB.PART 10.goat 1-good

But when they precede a number expression (which is an instance of QP in (50)), for example, they are separated by a pause, showing this to be a marked order:

(53) a. *mũ-ndũ ũ-mwe mũ-thĩĩnj-i mbũri*
1-person 1-one 1-slaughter-HAB.PART 10.goat
'one person who is a slaughterer of goats'

b. *mũ-ndũ mũ-thĩĩnj-i mbũri, ũ-mwe*
1-person 1-slaughter-HAB.PART 10.goat, 1-one

The agentive nominalizations clearly contrast in their word order possibilities, as we saw in (31). Compare, for example, the unmarked order of (51) with that of (54):

(54) *mũ-thĩĩnj-i mbũri w-a-kwa ũ-mwe*
1-slaughter-NOM 10.goat 1-ASSOC-my 1-one
'my one goat slaughterer'

In short, the agentive nominal occurs not in the position of an AP, following the head and other modifiers, but in the position of the head itself, preceding all other modifiers except for focused demonstratives.

We see, then, that the agentive nominalization has both the morphology and the syntactic positioning of the head of an NP. In these respects it behaves like a pure noun, which can never be used adnominally (without an associative particle). The agentive nominal head of the pure NP construction shares this pure nominal property:

(55) **mũ-ndũ mũ-thĩĩnj-i w-a mbũri*
1-person 1-slaughter-NOM 1-ASSOC 10.goat
Lit.: 'a person slaughterer of goats'

In sum, the mixed categories in question are both truly phrasal and truly mixed, in the sense that they consist of a VP embedded within an NP whose head position is occupied by the agentive nominalization.

3 Haspelmath's generalization

So far we have found evidence for the following two conclusions about the mixed agentive nominalization constructions in Gĩkũyũ:

(56) a. The agentive nominal heads of mixed categories in Gĩkũyũ are deverbal nouns occupying the phrase-initial head position of a nominal projection (NP).

b. These mixed category constructions in Gĩkũyũ consist syntactically of components of a verbal projection (VP) embedded within a nominal projection (NP).

(56a,b) have been established in Sections 1 and 2. We now observe that there is a relation between these two conclusions. The morphological structure of the head reflects the syntactic structure of the phrasal construction: the construction consists syntactically of a verbal phrase embedded within a nominal phrase, as we saw in (39) and (40), and the head contains a verbal base embedded within nominal morphology, as we see in (57). (The v, n subscripts indicate the categorial type of the stems as respectively verbal or nominal.)

(57) a. $[m\tilde{u}\text{-}[[th\tilde{i}\tilde{i}nj]_v\text{-}i]_n]_{\mathrm{N}}$
1-slaughter-NOM

b. $[m\tilde{u}\text{-}[[in\text{-}\tilde{i}r]_v\text{-}i]_n]_{\mathrm{N}}$
1-sing-APPLIC-NOM

This relationship is not to be dismissed as an accident or a purely language-particular phenomenon. The existence of similar morphology-syntax relations in mixed categories is widespread crosslinguistically, and has been generalized by Haspelmath 1995, who specifically relates the syntactic structure of a mixed category to the morphological structure of the head (Haspelmath 1995: 56–58):

(58) **Haspelmath's generalization:**

(a) In words derived by *inflectional* word-class-changing morphology, the internal syntax of the base tends to be preserved.

(b) In words derived by *derivational* word-class-changing morphology, the internal syntax of the base tends to be altered and assimilated to the internal syntax of primitive members of the derived word-class.

Haspelmath defines 'inflectional' morphology as productive morphology. By 'internal syntax' Haspelmath refers to the combination of the head with its dependents inside its phrase; the 'external syntax'—how the head combines with elements outside its phrase—is determined by the derived word class (Haspelmath 1995: 52).

Nikitina (2005, 2006) shows that constructions with mixed-category syntax occur in some Mande languages which lack formal word-class changing morphology. Thus the morphology-syntax relationship formulated by Haspelmath is only a one-way implication: productive word-class changing morphology is associated with mixed-category syntax, but mixed-category syntax can also arise independently. With this understanding of its limitations, we reformulate the generalization in our terms as in (59):

(59) The productive morphological derivation of a word of one category $\mathcal{C}_1$ from a base of another category $\mathcal{C}_2$ will tend to preserve the syntactic structure of $\mathcal{CP}_2$ within the syntactic context of $\mathcal{CP}_1$, while less productive category-changing morphology will tend to alter the syntactic context of the base category $\mathcal{CP}_2$ to that of $\mathcal{CP}_1$.

For Gĩkũyũ $\mathcal{C}_1 = \text{N}$ and $\mathcal{C}_2 = \text{V}$. Thus, the agentive nominalization is a nominal word of category N productively derived from a verbal base of category V (the verb stem), and VP structure is preserved within the syntactic context of NP.

4 Implications for theories of mixed categories

Haspelmath's generalization and its particular instantiation in Gĩkũyũ are highly problematic for one previous approach to mixed categories, which we call 'the single projection' hypothesis:

(60) **The single-projection hypothesis:**
A mixed category is the single phrasal projection of a morphologically 'mixed' (underspecified, indeterminate, bivalent) head.

In precisely what way the head is morphologically 'mixed' under the single-projection hypothesis (60) varies with the particular version of the approach. The feature-neutralization version assumes that the head of the mixed construction is lexically underspecified for its category and so projects a categorially indeterminate phrasal structure which may contain constituents of mixed

types (Aoun 1981, van Riemsdijk 1983, Grimshaw 1991). A consequence of this approach is that mixed category constructions must have underspecified heads which are formally ambiguous as to category type—just as we see in English gerundive verb constructions (2) or the Italian *infinito sostantivato* (5), (6), (8). In contrast, a type-hierarchical version of the single projection hypothesis assumes that such mixed categories belong to a distinct fine-grained category that inherits some typical properties from nouns and some from verbs (Malouf 1998: 89, 163). Under the former (feature-neutralization) version the mixed category is thus underspecified or neutral in category, while on the latter (type-hierarchical) version it is multiply-specified or bivalent. Under both versions of this approach it heads a single endocentric projection and the lexical integrity of the head is preserved.

A basic problem for the feature-neutralization version of (60) is that category neutrality of the head is *not* a universal characteristic of mixed category constructions, as Gĩkũyũ shows. Categorially unambiguous heads also appear in Quechua nominalization-headed clauses (Lefebvre and Muysken 1988), Arabic deverbal process nominals or *maṣdars* (Fassi Fehri 1993), Hebrew action nominalizations (Hazout 1995, Falk 2006), and many other examples (Haspelmath 1995).

A second problem, which applies to all varieties of the single projection hypothesis, is that phrasal coherence constrains the mixing of categories. That is, mixed category constructions (in configurational languages, at least) do not freely mix or interleave constituents of the different category types, but instead cohere within distinct regions which can be bounded by distinct phrase structure brackets. For example, the VP-style constituents within the Gĩkũyũ mixed category cluster together as a unit, preventing higher NP-style elements from interrupting them. We have observed this property in Gĩkũyũ in Sections 1.3 and 2.1, and represented it by the tree structure (39).

Phrasal coherence appears to be a general property of mixed category constructions across languages (Bresnan 1997). With the Italian *infinito sostantivato*, for example, the constituents preceding the infinitive are always nominal (determiners and adjectives) and can cooccur with either post-infinitive VP constituents (such as direct objects and adverbs) or NP constituents (such as postnominal adjectives and *di* phrases). However, the post-infinitive constituents of different category types cannot cooccur, but must be uniformly of the VP type or the NP type. As shown in the following examples from Zucchi (1993: 222) and Bresnan 1997, a post-infinitive adjective permits only nominal constituents (e.g. other adjectives and *di* phrases) to follow, while a

post-infinitive adverb permits only verbal constituents (e.g. a direct object or other verbal complement) to follow.

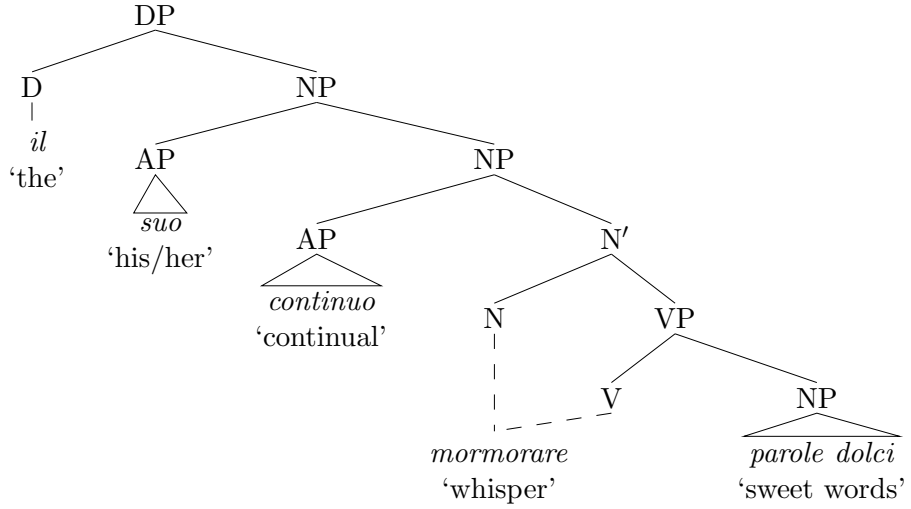
(61) a. *il mormorare sommesso/*sommessamente del mare*
the whisper.INF soft/softly of.the sea
'the soft whispering of the sea' (Zucchi 1993: 220)

b. *il suo mormorare sommessamente*
the his/her whisper.INF softly
'his/her whispering softly' (Zucchi 1993: 226)

c. *il suo momorare continuamente/*continuo parole dolci*
the his/her whisper.INF continually/continual sweet words
[compare to (5)] (Zucchi 1993: 245)

This phrasal organization suggests that the infinitival head may take a VP complement, which prevents a postnominal adjective (required to appear in postnominal position adjacent to the head) from appearing. Bresnan 1997 depicts the syntactic structure of the Italian mixed category construction in the following diagram:[17]

(62) Italian *infinito sostantivato*:



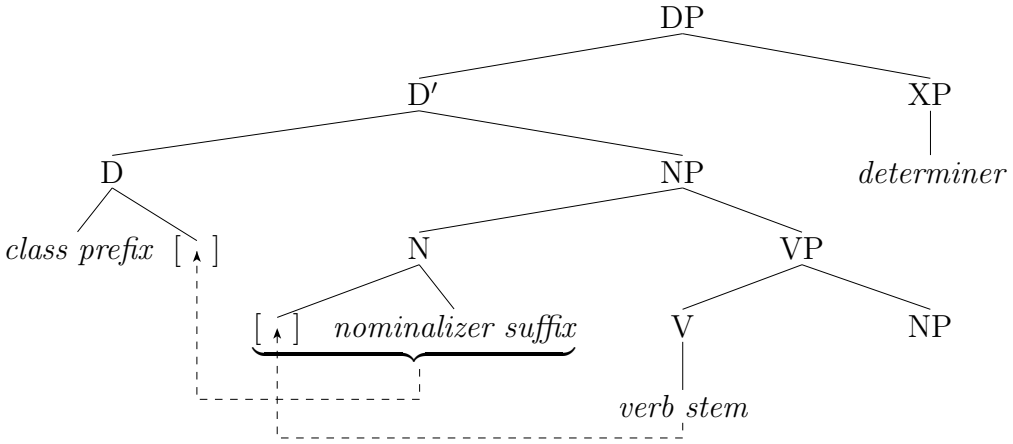
These considerations motivate the dual-projection hypothesis:

[17]Zucchi (1993: 251) analyzes the infinitive as dominated by V, but all of the evidence he cites is consistent with its being dominated by N.

(63) **The dual-projection hypothesis:**
Mixed categories consist of not one, but two projections that differ in category type in a way reflected in the morphology of the head.

The most widely known version of the dual-projection hypothesis assumes that the verbal base of a deverbal nominal mixed category starts out as a verb heading the VP and then is moved into the N position (or to the position of a nominal functional projection), as illustrated in (64) (cf. Fassi Fehri 1993, Hazout 1995, Borsley and Kornfilt 2000). It is schematically applied to the Gĩkũyũ construction in (64):

(64) Syntactic word-formation by head movement:



By further assuming that the derived X^0 category is moved into the next higher X^0 category, this theory can also explain Haspelmath's generalization (59): the morphemic structure of the agentive nominalization, under these assumptions, must reflect the syntactic embedding relations of the projections.

This approach to mixed categories preserves the principle of endocentricity, explaining how two different categories of syntactic projections can arise from a single word: the categories are separately projected from different heads which are subsequently joined by syntactic movement into a single word, and it captures the systematic relation between the morphological composition of the head and the syntactic structure of the mixed category.

The weakness of the approach is in failing to explain the relations between lexically and syntactically derived words. Lexically derived words do not give evidence of phrasal sources for their morphological components. This point is made by Ackema and Neeleman (2001) for English, using the following exam-

ple:[18]

(65) *the truck driver reluctantly a rusty to Arizona
'the one who reluctantly drives a rusty truck to Arizona'

If the synthetic compound *truck driver* were syntactically derived from phrasal syntactic sources by head-movement, then the specifiers, complements, and modifiers of the phrasal sources of the morphological components should be visible, exactly as is hypothesized in the case of the mixed category construction (64). This point is illustrated by (66):

(66) the -er [$_{VP}$ reluctantly drive a rusty truck to Arizona] ⇒
the -er [$_{VP}$ reluctantly truck-drive a rusty __ to Arizona] ⇒
the truck drive-er [$_{VP}$ reluctantly __ a rusty __ to Arizona]

In Gĩkũyũ, too, there is evidence that agentive nominals heading the pure NP constructions cannot be syntactically derived. If these were derived in the syntax in the way hypothesized for the heads of the mixed NP/VP constructions (64), then we would expect to find evidence of a syntactic VP, such as stranded adverbials or the like, even in the unmixed NP construction. But nothing of this sort can appear:

(67) a. *mũ-in-i wega w-a nyĩmbo
1-sing-NOM well 1-ASSOC 10.song
Lit.: 'a singer well of songs'

b. *mũ-in-i w-a nyĩmbo wega
1-sing-NOM 1-ASSOC 10.song well
Lit.: 'a singer of songs well'

In (67) the selection of the associative phrase *wa nyĩmbo* to express the nominal complement (perhaps as an argument adjunct) clearly marks the construction as unmixed. This inference follows from the uniformity of selection of complement type, described as 'lexical coherence' in Section 1.3. As such, it cannot take an adverbial modifier. Contrast the construction in (68), where the adjunct 'of the settlement' is not selected by the nominalization and hence is

[18]The same point is made by Bresnan and Mchombo 1995 for Bantu noun class prefixal morphology.

consistent with the structure of a mixed NP/VP construction. Here an adverb is possible (68a), but only in the coherent VP portion of the structure, which precedes the higher adnominal adjunct (68b):

(68) a. *mũ-in-i* *wega* *w-a* *i-tũũra*
1-sing-NOM well 1-ASSOC 5-settlement
Lit.: 'a singer well of the settlement'

b. **mũ-in-i* *w-a* *i-tũũra* *wega*
1-sing-NOM 1-ASSOC 5-settlement well
Lit.: 'a singer of the settlement well'

The inability of the pure NP agentive nominals to take manner adverbs holds for both Gĩkũyũ and English, as the literal translations of (67a,b) show. Rappaport Hovav and Levin 1992 propose an account of why English *-er* nominals, although showing some eventive properties, nevertheless prohibit certain types of adverbs. They suggest that such adverbs are modifiers of an open event variable in the argument structure of a base verb, but that in the case of agentive nominals, this event variable is lexically quantified prior to syntactic argument linking (Rappaport Hovav and Levin 1992: 143). On this account, the possibility or impossibility of such adverbial modifiers follows from a *lexical* property of the nominalized forms, and is expressed in their lexical argument structures.

Given that there are lexically derived agentive nominals, the problem for syntactic word formation is twofold. First, words hypothesized to be syntactically derived do not differ in morphological structure from those lexically formed (see Bresnan and Mchombo 1995 for a review of evidence). While this fact can be captured by various stipulations, it remains fundamentally unexplained by the syntactic word-formation approach, because the opposite state of affairs could be captured just as readily and would in fact be seen as confirmation of the theory. Second, the question of which words are lexically and which syntactically derived—or to put it more neutrally, which words head unmixed and which head mixed category constructions—needs to be answered by the syntactic word formation approach just as much as by other approaches.

5 An analysis within LFG

Within LFG there is a simple solution to these problems posed by Gĩkũyũ. Suppose that each lexeme carries a categorization constraint which is preserved

under productive morphological processes. Such a constraint is easily formalized via inside-out function application using the 'Cat' function (Halvorsen and Kaplan 1988; Nordlinger 1998; Kaplan 1995; Crouch, Dalrymple, Kaplan, King, Maxwell, and Newman 2006). For example, a verb lexeme would carry a constraint like that in (69a), and a noun lexeme would carry one like that in (69b):

(69) a. VP ε Cat((PRED$\uparrow$))

b. NP ε Cat((PRED$\uparrow$))

Such constraints categorize the c-structure domain in which a lexical head (providing the PRED attribute) must be found. Technically, the constraints require that a VP (respectively NP) be among the c-structure categories of the nodes in the inverse image of the ϕ mapping from the f-structure containing the PRED.

Productive morphological processes such as tense-marking or number inflection will preserve categorization constraints. If the English verb *slaughter*, for example, carries the constraint (69a), so will its present tense form *slaughters*. In contrast, derivational morphology usually does not preserve the categorization information of the base lexeme. For example, the argument structure of the English deverbal noun *slaughterer* is derived from its verbal base *slaughter*. The lexical relation of *slaughterer* to *slaughter* is relatively transparent, as illustrated in (70):[19]

(70) slaughter: 'slaughter$< x, y >_v$'
slaughterer: 'agent-of $< x$, slaughter$< x, y >>_n$'

The notations '$\langle \ldots \rangle_v$' and '$\langle \ldots \rangle_n$' represent the categorization of the predicators as verbal or nominal, respectively. Note that the categorization of the base verb is not retained in the nominalization of the verb. These features of the argument structures will flag the presence of the categorization constraints in (69) in the lexical entries for these predicators (which are presumably derived by some version of the lexical mapping theory):[20]

[19]Recall that extracting the agent role is only the most typical function of the agentive nominalizing suffix, as mentioned at the outset of Section 1.

[20]The lexical entry forms show only the grammatical functions required for completeness and coherence, abstracting away from the argument relations among base and derivative shown in (70).

(71) a. slaughter: V: (↑ PRED) = 'slaughter<(↑ SUBJ)(↑ OBJ)>$_v$'
v: VP ε Cat((PRED↑))

b. slaughterer: N: (↑ PRED) = 'slaughterer<(↑ OBL$_\theta$)>$_n$'
n: NP ε Cat((PRED↑))

For Gĩkũyũ we simply assume that mixed categories are productively formed words which retain the categorization constraints of their bases, as in (72):

(72) mũthĩĩnji: 'agent-of $< x$, slaughter$< x, y >_v >_n$'

The argument structure of the Gĩkũyũ agentive nominalization in (72), unlike the English (70), has the categorization information '$\langle \ldots \rangle_v$' embedded within it.

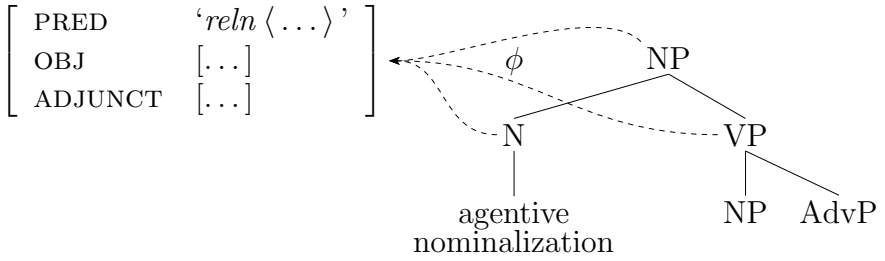
Predicators of the type in (72) are formed in the component of grammar which produces argument structures, in this case the lexical morphology. The verbal argument structure is transparently embedded within a nominally categorizing argument structure (contributed by the nominalizing morphology and designated '$\langle \ldots \rangle_n$'). From this information a lexical entry such as (73) can be derived:[21]

(73) mũthĩĩnji: N: (↑ PRED) = 'slaughterer<<(↑ OBJ)>$_v$>$_n$'
v: VP ε Cat((PRED↑))
n: NP ε Cat((PRED↑))

This analysis permits a complete and coherent f-structure for the entire construction. To see this, consider the following. The above lexical entry (73) requires that the f-structure of the PRED must be the image under ϕ of VP as well as NP. In other words, the lexical entry licenses the presence of both a VP and an NP. The head N contributes to f-structure both the noun class required of every NP in Gĩkũyũ and the attributes of predicator, while the VP allows an object and adverbial adjunct, which are characteristic of VP f-structures, as illustrated in (74).

[21] Again, the lexical entry forms show only the grammatical functions required for completeness and coherence, abstracting away from the argument relations among base and derivative shown in (72).

(74)



The question arises, given (73), Why must VP be inside NP rather than the other way around? An answer is provided by extended head theory (Jar n.d.; Zaenen and Kaplan 1995: 221–2; Bresnan 2001): an extended head by definition cannot appear lower in the tree than the phrase(s) which it heads. Hence the nominalization's NP projection must dominate the VP.[22]

Of course, not every nominalized verb will be able to serve simultaneously as a VP and NP predicator. In Gĩkũyũ, we find that agentive and other nominalizations can head mixed categories, while infinitive nouns cannot (Mugane 1996). In Italian it is the reverse (Zucchi 1993). In English, agentive nominalizations are unmixed nominals, which take only nominal complements and modifiers (75a), and reject the objects and adverbs of verbal constructions (75b,c):

(75) a. this unthinking slaughterer of goats

b. *this slaughterer goats unthinkingly

c. Don't slaughter goats unthinkingly!

As illustrated above (71), the PRED value of the English nominalization simply lacks a transparently embedded verbally categorizing argument structure corresponding to its verbal base.

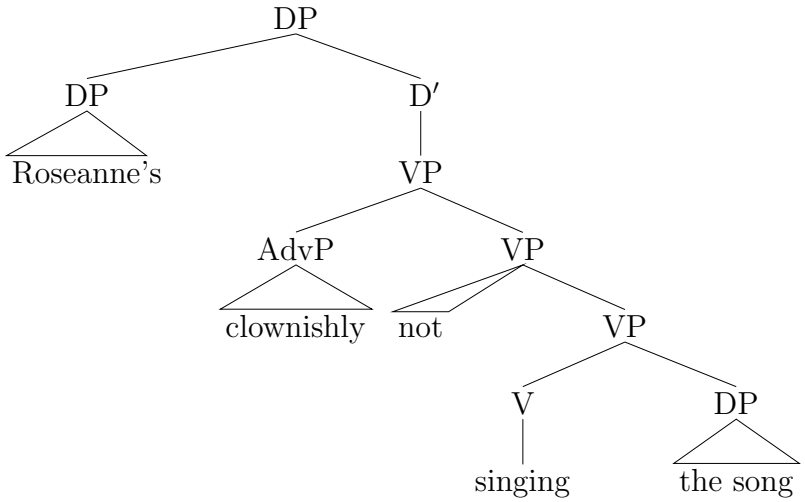
Thus, the lexical morphology of a language must provide the resources to support mixed categories in the syntax by licensing appropriate f-structure attributes. But argument structure alone will not suffice to solve the syntactic

[22]To see this, note that in (74) the agentive nominalization is a noun and is the c-structure head (as well as the extended head) of the NP which dominates it. It is also the extended head of its VP sister, which is annotated by the principle permitting lexical categories to have co-heads as an option (Bresnan 2001: ch. 6). Hence in (74) the f-structures of the N and VP are identified through unification as permitted by the extended head theory.

problems of phrasal coherence, endocentricity, and head positioning presented by mixed categories. For these, the theory of structure-function mapping appears essential.

The extended head theory of mixed categories makes an interesting prediction about the syntactic positioning of the heads in their phrasal structures. Suppose that a mixed category involves a lexical category such as VP embedded in a functional category such as DP. In this case, the head may be positioned in VP without violating endocentricity. Every lexical category must have an extended head, but a functional category need not, because functional categories are headed by recoverable classes of elements (Bresnan 2001: ch. 7). This gives us a natural structure for the English gerundive construction (Bresnan 2001: ch. 13) illustrated in (76).

(76)



The Gĩkũyũ-style analysis given above (74) would be inappropriate for the verbal English construction because it would have the gerundive verb in the head N position of the mixed category, where we would expect the possibility of prenominal AP modifiers (such as are found in the Italian infinitive noun construction) and nominal negative prefixation cooccurring with the VP properties:

(77) Roseanne's clownish non-singing *(of) the national anthem.

There is also positive evidence for the presence of DP in (76). The DP explains why no nominal head is required in the construction: the nominal functional category DP need not have a head. The DP also explains why the

subject of the verbal gerundive construction takes the genitive form, because this is the syntactic attribute of the specifier of DP.[23]

There is in fact interesting evidence from quantifier scope that the genitive NP has the scope properties of possessive NPs of nouns, and not of subjects of embedded Ss. Observe the contrast in (78) (Zucchi (1993: p. 50):

(78) a. John resents everyone's taking a day off.

b. John resents that everyone takes a day off.

The quantifier phrase in (78a) may have wide scope, exactly as in (79):

(79) John resents everyone's absence.

Both (78a) and (79) are ambiguous: John may resent only the universal absence of other employees, leaving him stuck with all the work (the wide scope reading); or for each absent employee, John may resent that person's individual absence (the narrow scope reading). But (78b) differs in preferring the narrow scope reading.

Finally, in the context of the theory of extended heads the DP in (76) can explain why the verbal gerund shares some properties of deverbal nouns: a nominally categorizing argument structure is needed to support a possessor. Thus the two types of gerundive verb forms in English can be represented in our theory as in (80a,b):

(80) a. singing: N: ($\uparrow$ PRED) = 'singing$<(\uparrow \text{OBL}_\theta)>_n$'

b. singing: V: ($\uparrow$ PRED) = 'singing$<<(\uparrow \text{SUBJ})(\uparrow \text{OBJ})>_v>_n$

Note that the outer nominal typing of the verbal argument structure in (80b) does not prevent the categorization of the gerundive as a V in c-structure. This is evidence that the category identity properties of heads can be distinct from their f-structure licensing properties, as we have assumed. If, however, the verbal gerund in (80b) were categorized as N instead of V, a mixed lexical category construction would result, allowing examples such as (77). Interestingly, examples of this type did occur in historically earlier stages of English (Tajima 1985).

[23]We may assume that the Specifier of DP is the most prominent argument function for verbs or nouns: POSS or SUBJ, depending on the a-structure requirements (Laczkó 1995, 1997).

We see, then, that when category mixing involves a lexical and a functional category, the head may appear in the lower, lexical category. But when category mixing involves two lexical categories sharing the same head, it is predicted that the head must appear in the upper lexical category. This follows because by endocentricity every lexical category must have an extended head, and an extended head by definition cannot appear lower in the tree than the phrase(s) which it heads.

Finally, we observe a limitation of our solution. We have reconciled the conflict between the principles of endocentricity and lexical integrity by exploiting the fact that words and phrases talk to each other through their common functional structure. Thus a single lexical word such as a denominal agentive nominalization can constrain the category types of the regions of tree structure that correspond to its functional domain. At the same time, a single lexical head in constituent structure can serve as the extended head of a cascade of phrases in the tree structure above or below it through the many-to-one correspondence of tree structure nodes to functional structures. This mapping between expressions and functional structure is intentionally imperfect: by flattening trees, it loses information. This property of the correspondence architecture is considered a feature, not a bug, because many languages in fact make far less use of hierarchical constituent structure than do highly endocentric languages like English (Bresnan 2001 and references). However, it follows from this property that only a minimal amount of matching between the word-structure and the constituent structure can be explained by the analysis offered here. To extend the matching between more than two levels of morphological derivation and syntactic tree structure would require that word derivation define hierarchical functional structures (Simpson 1991, Nordlinger 1998), but that is beyond the scope of the present study.

6 Concluding Note

The analysis of mixed categories we have presented draws heavily on the flexibility and power of the LFG architecture, and in particular on the central conception of the ϕ mapping between categorial structures and feature structures (as well as some of its specific applications)—which are due to Ron Kaplan. Thanks, Ron!

References

Ackema, Peter and Ad Neeleman. 2001. Competition between syntax and morphology. In *Optimality Theoretic Syntax*, ed. by Geraldine Legendre, Jane Grimshaw, and Sten Vikner, 29–60. Cambridge, Mass.: MIT Press.

Alsina, Alex. 1996. *The Role of Argument Structure in Grammar. Evidence from Romance*. Stanford: CSLI Publications.

Andrews, Avery D. 1990. Unification and morphological blocking. *Natural Language & Linguistic Theory* 8.507–57.

Aoun, Yosef. 1981. Parts of speech: a case of redistribution. In *Theory of Markedness in Generative Grammar*, edited by Adriana Belletti, Luciana Brandi, and Luigi Rizzi, 3–24. Pisa: Scuola Normale Superiore di Pisa.

Austin, Peter and Joan Bresnan. 1996. Non-configurationality in Australian aboriginal languages. *Natural Language & Linguistic Theory* 14.215–68.

Borsley, Robert D. and Jaklin Kornfilt. 2000. Mixed extended projections. In *The Nature and Function of Syntactic Categories*, ed. by R. D. Borsley. New York: Academic Press, 101–131.

Bresnan, Joan. 1997. Mixed categories as head-sharing constructions. *Proceedings of the LFG97 Conference, University of California, San Diego*, ed. by Miriam Butt and Tracy Holloway King. On-line, CSLI Publications: http://www-csli.stanford.edu/publications/.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan and Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language* 63.741–82.

Bresnan, Joan and Samuel A. Mchombo. 1995. The Lexical Integrity Principle: Evidence from Bantu. *Natural Language & Linguistic Theory* 13.181–252.

Crouch, Richard, Mary Dalrymple, Ronald M. Kaplan, Tracy H. King, John T. Maxwell III, and Paula Newman. 2006. *XLE Documentation*. On-line, Palo Alto Research Center: www2.parc.com/istl/groups/nltt/xle/doc/xle_toc.html.

Falk, Yehuda. 2006. Constituent structure and grammatical functions in the Hebrew nominal phrase. To appear in *Architectures, Rules and Preferences*,

ed. by Jane Grimshaw, Joan Maling, Chris Manning, Jane Simpson and Annie Zaenen. Stanford: CSLI Publications.

Fassi Fehri, Abdelkader. 1993. *Issues in the Structure of Arabic Clauses and Words*. Dordrecht: Kluwer.

Grimshaw, Jane. 1991. Extended projection. Waltham, Massachusetts: Brandeis University Linguistics and Cognitive Science Program MS.

Halvorsen, Per-Kristian and Ronald M. Kaplan. 1988. Projections and semantic description in Lexical-Functional Grammar. *Proceedings of the International Conference on Fifth Generation Computer Systems, 1988, Tokyo, Japan*, 1116–1122.

Haspelmath, Martin. 1995. Word-class-changing inflection and morphological theory. In *Yearbook of Morphology 1995*, ed. by Geert Booij and Jaap van Marle, 43–66. Dordrecht: Kluwer Academic Publishers.

Hazout, Ilan. 1995. Action nominalizations and the Lexicalist Hypothesis. *Natural Language & Linguistic Theory* 13.355–404.

Jar, M. [John Maxwell III, Annie Zaenen, Ronald M. Kaplan, and Mary Dalrymple]. n.d. Reconstituted X′ constituents in LFG. Palo Alto, CA: Xerox Palo Alto Research Center duplicated MS.

Kaplan, Ronald M. 1995. The formal architecture of Lexical-Functional Grammar. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, eds., *Formal Issues in Lexical-Functional Grammar*. Stanford: CSLI Publications.

Laczkó, Tibor. 1995. *The Syntax of Hungarian Noun Phrases — A Lexical-Functional Approach*. Frankfurt am Main: Peter Lang.

Laczkó, Tibor. 1997. Action nominalization and the possessor function. *Acta Linguistica Hungarica* 44.3–4: 413–475

Lee, Hanjung. 1999a. The domain of grammatical case in Lexical-Functional Grammar. *Proceedings of the LFG99 Conference, University of Manchester*, ed. by Miriam Butt and Tracy Holloway King. On-line, CSLI Publications: http://www-csli.stanford.edu/publications/.

Lee, Hanjung. 1999b. Aspectual and Thematic Licensing of Grammatical Case. *CLS* 35: 203–222.

Lefebvre, Claire and Pieter Muysken. 1988. *Mixed Categories: Nominalizations in Quechua*. Dordrecht: Kluwer.

Malouf, Robert P. 1998. *Mixed Categories in the Hierarchical Lexicon*. Stanford: Stanford University Department of Linguistics Ph.D. dissertation.

Malouf, Robert P. 2000. *Mixed Categories in the Hierarchical Lexicon*. Stanford: CSLI Publications.

Mchombo, Sam A. 1978. *A Critical Appraisal of the Place of Derivational Morphology within transformational Grammar, Considered with Primary Reference to Chichewa and Swahili,* London: School of Oriental and African Studies, University of London Ph.D. dissertation.

Mugane, John. 1996. *Bantu Nominalization Structures.* Tucson, Arizona: University of Arizona Department of Linguistics Ph.D. dissertation.

Mugane, John. 1997. *A Paradigmatic Grammar of Gĩkũyũ*. Stanford Monographs on African Languages. Stanford: CSLI.

Mugane, John. 2003. Hybrid constructions in Gĩkũyũ: Agentive nominalizations and infinitve-gerund constructions. In *Nominals Inside and Out*, edited by Miriam Butt and Tracy Holloway King, 235–265.

Myers, Scott P. 1987. *Tone and the Structure of Words in Shona*. Amherst: University of Massachusetts, Amherst, Ph.D. dissertation.

Nikitina, Tatiana. 2005. Mixed category constructions and word order change in Niger-Congo. Paper presented at the XVIIth International Conference on Historical Linguistics: Workshop on Constructions in Language Change (University of Wisconsin, Madison, July 31–August 5, 2005.

Nikitina, Tatiana. 2006. Morphological exponent of mixed category constructions: Embedded clauses with mixed syntactic properties in Wan. Stanford: Stanford University Department of Linguistics MS.. Submitted for publication.

Nordlinger, Rachel. 1998. *Construction Case: Evidence from Australian Languages*. Stanford: CSLI Publications.

Przepiórkowski, Adam. 1999. On case assignment and "adjuncts as complements". In *Lexical and Constructional Aspects of Linguistic Explanation*, ed. by Gert Webelhuth, Jean-Pierre Koenig and Andreas Kathol, 231–245. Stanford: CSLI Publications.

Rappaport Hovav, Malka and Beth Levin. 1992. *-er* nominals: implications for the theory of argument structure. In *Syntax and the Lexicon. Syntax*

and Semantics. Vol. 26, ed. by Tim Stowell and Eric Wehrli, 127–153. San Diego: Academic Press.

van Riemsdijk, Henk. 1983. The case of German adjectives. In *Linguistic Categories: Auxiliaries and Related Puzzles*, ed. by Frank Heny and Barry Richards, 223–52. Dordrecht: D. Reidel.

Simpson, Jane. 1991. *Warlpiri Morpho-Syntax: A Lexicalist Approach.* Dordrecht: Kluwer.

Tajima, Matsuji. 1985. *The Syntactic Development of the Gerund in Middle English.* Tokyo: Nan'un-do.

Wechsler, Stephen and Yae-Sheik Lee. 1996. The domain of direct case assignment. *Natural Language & Linguistic Theory* 14: 629–664.

Zaenen, Annie and Ronald M. Kaplan. 1995. Formal devices for linguistic generalizations: West Germanic word order in LFG. In *Formal Issues in Lexical-Functional Grammar*, ed. by Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, 215–39. Stanford, California: CSLI Publications.

Zucchi, Alessandro. 1993. *The Language of Propositions and Events.* Dordrecht: Kluwer Academic Publishers.