

# THE STRUCTURAL ORIGINS OF TONAL OVERLAYS IN TOMMO SO (DOGON) COMPOUNDS

**Laura McPherson**

UCLA, Los Angeles

[laura.emcpherson@gmail.com](mailto:laura.emcpherson@gmail.com)

Tommo So (Dogon, Mali) has two classes of root compounds, distinguished from one another by tonal behavior. In the first class, canonical compounds, all non-final stems have their tone replaced by L in a process I call tone lowering. This is the same process seen when a noun is modified by certain grammatical categories (adjective, demonstrative, relative clause), but I show that this is a surface coincidence, with compound tone lowering driven by phonotactic considerations and modifier tone structurally determined. The other class of compounds, pseudo-genitive compounds, lowers the tone of the final stem(s) in a pattern of tone lowering identical to true genitive constructions. I argue that in this case, the compounds are lexicalized syntactic phrases, based on evidence such as headedness and the ability to contain regular inflection, and that given this structure, tone lowering in pseudo-genitive compounds and true genitives is one and the same.

Tommo So (Dogon, Mali) a deux classes de mots composés, qui se distinguent l'une de l'autre par leurs propriétés tonales. Dans la première classe, celle des composés canoniques, tout lexème non-final subit un abaissement tonal dans lequel le ton du lexème est remplacé par un ton bas. On observe ce même processus quand un nom est modifié par certaines catégories grammaticales (adjectif, démonstratif, proposition relative), mais il est démontré dans cet article que cette similitude n'est qu'une coïncidence superficielle ; l'abaissement tonal du composé est entraîné par des propriétés phonotactiques, alors que le ton des syntagmes modifiés suit de propriétés structurales. L'autre classe de mots composés, les composés pseudo-génitifs, impose l'abaissement tonal sur le lexème final (ou les lexèmes finaux), dans un système d'abaissement tonal identique à celui des vrais génitifs. Sur la base de la position du mot tête de série et de la capacité du composé à contenir de la morphologie flexionnelle régulière, l'argument est que, dans ce cas, les composés sont en fait des syntagmes lexicalisés, et qu'étant donné cette structure, l'abaissement tonal dans les composés pseudo-génitifs et dans les vrais génitifs est le même.

## 0. INTRODUCTION

The tonal system of Tommo So, a Dogon language of Mali, is characterized by heavy use of tonal overlays that completely erase a stem's lexical tone. One area in which these overlays are found is nominal compounds, which make up a large portion of the language's nominal vocabulary. Out of 4,405 nominal entries in the lexical corpus, 1,271 of those are compounds. These 1,271 compounds can be further subdivided into root compounds (those with a nominal head) and synthetic compounds (those with a deverbal head). In the interest of space, this paper focuses solely on root compounds, which fall into two classes distinguished by tonal overlays. The first group of compounds I call "canonical compounds". They are more common than the second group, "pseudo-genitive compounds", and are also more diverse in terms of headedness. I will argue that the tonal differences between the two groups are reflexes of underlying structural differences, with the canonical compound tone indicating the traditional branching N structure proposed for compounds, while pseudo-genitive tone indicates the lexicalized use of a possessive structure. Interestingly, the possessive structure for compounds is not the same as what the true possessive structure would be for the stems involved; instead, all nouns, alienable or inalienable, are treated as inalienable when it comes to compounds. Further, I argue that the proposed structures correctly predict the asymmetries in headedness seen in the two classes of compounds.

In section 1, I give a brief introduction to the Tommo So language. In section 2, I turn to canonical compounds, arguing that their tonal overlays derive from the fact that

they are syntactically a branching N node and, as such, they are subject to the same tonotactic constraints as simplex nouns. A surface similarity in tone with N+Adj constructions is shown not to reflect an underlying structural similarity. Section 3 investigates pseudo-genitive compounds, which, as the name suggests, have a structure parallel to that of possessive constructions in the language. Their tone is a consequence of their phrasal structure, which also allows limited cases of regular inflection inside of these compounds. Section 4 shows how some compounds vary in their tonal pattern and how this phenomenon is driven by variation in whether the modifying stem is head-adjoined to the head of the compound or is placed in the specifier of its NP (possessor position). Finally, section 5 draws conclusions and suggests directions for future research.

### 1. TOMMO SO<sup>1</sup>

Tommo So (McPherson, forthcoming) is spoken by around 60,000 people on the rocky plateau stretching between Bandiagara and Douentza in Mali's Mopti Region. Prior to the current fieldwork, the only documentation of the language was Plungian (1995), a short grammatical sketch laying out the basics of the morphology. Unfortunately, this sketch failed to mark tone, which plays a very important grammatical role in Tommo So, as it does in all Dogon languages (Heath & McPherson, forthcoming). In what follows, we will see one such application of grammatical tone overlays.

Tommo So is a two-tone language, and the only licit stem melodies for native words are /H/ and /LH/.<sup>2</sup> For example:

(1)	/H/		/LH/	
	<b>náá</b>	'mother'	<b>nàá</b>	'cow'
	<b>dámmá</b>	'village'	<b>dàmmá</b>	'type of hoe'
	<b>kégéré</b>	'saddle'	<b>kèbèlé</b>	'shard'

These tonotactics will become important later in the discussion of canonical compounds. In terms of morphology, it is a somewhat split language, with the nominal domain displaying isolating morphology<sup>3</sup> (characterized by stem-level tonal overlays and clitics) and the verbal domain displaying agglutinating morphology (with both derivational and inflectional suffixes alongside tone changes).

Given these two lexical tonal melodies (with a few exceptional cases of /HL/ introduced from loanwords), we would expect to see the following combinations of tones resulting from the concatenation of stems into compounds: H-H, H-LH, LH-LH, LH-H. As I will demonstrate below, none of these is attested. In every case, all but one stem in the compound has its tone overwritten with {L}, an impossible tonal melody for stems in isolation.

<sup>1</sup> The transcription system used here is more or less IPA with the following deviations: /r/ refers to an alveolar tap, /y/ to a palatal glide, and /j/ to a voiced alveopalatal affricate. Vowel length is represented by doubling the vowel, and nasalization is marked with a superscript 'n'. Hyphens indicate affix boundaries and equals signs (=) indicate clitic boundaries.

<sup>2</sup> Tone melodies in slashes indicate a stem-level lexical melody, which is mapped onto any number of available moras, though not in an automatic fashion. That is, a melody /LH/ can be realized on a trisyllabic stem as either LLH or LHH, as in **kèbèlé** 'shard' and **kògódó** 'shell'. This is contrast to melodies in curly brackets (e.g. {L}), which indicate grammatical overlays.

<sup>3</sup> With the exception of the restricted human singular and plural suffixes, **-ne** and **-m**, and a diminutive suffix **-ý**.

## 2. CANONICAL COMPOUNDS

Canonical compounds (henceforth CC) are so-called because they are the most common type of root compound. They are also the most diverse in terms of their headedness, with left-headed and externally-headed compounds attested alongside the more common right-headed compounds. Like all compounds in Tommo So, there are no segmental changes that link one stem to another. They are simply created from direct juxtaposition of two or more stems accompanied by the tonal changes discussed below.

### 2.1 TONAL SHAPE

We can schematize the tonal shape of CCs as  $X^L Y$ . That is, X (representing any non-final stem in the compound) does not surface with its lexical {H} or {LH} melody, but instead surfaces with a {L} overlay (“tone lowering” [Heath & McPherson, forthcoming]), while Y (the final stem) retains its lexical tone. It is clear that this is not simply a case of tonal deletion, since Tommo So has a surface three-way contrast between H, L and  $\emptyset$ , with the pitch of  $\emptyset$  derived by interpolation between surrounding tones; see McPherson (2011) for further discussion. In the case of compounds, however, the tone of non-final stems is realized as a flat stretch of L. For example:

- (2) a. **tènné** + **kòró** → **tènnè<sup>L</sup>** **kòró**  
 well calabash well calabash  
 ‘calabash for drawing water’
- b. **màlbá** + **dúmmó** → **màlbà<sup>L</sup>** **dúmmó**  
 gun butt gun butt  
 ‘butt end of a rifle’

In the examples above, superscript L indicates that tone lowering has taken place. Compounds can be recursive, and in this case, only the final stem retains its lexical tone. Thus, if we add the stem **pógúru**<sup>4</sup> ‘belt’ to the compound **màlbà<sup>L</sup> dúmmó** ‘butt end of a rifle’, we get **màlbà<sup>L</sup> dùmmò<sup>L</sup> pógúru** ‘leather covering on the butt end of a rifle’, with tone lowering now operative on **dúmmó**; only **pógúru** retains its lexical /H/. Given the prevalence of floating associative tones in West African languages (e.g. Nissim 1977 on East Bamileke, Cahill 2007 on Konni, Williamson 1986 on Igbo, and Schuh 2009 on Ngamo, among others), we may ask whether this system of {L} overlays can be analyzed as a floating L that spreads leftward to replace the tone of the first stem. As we will see below, possessive constructions also involve tone lowering, but in this case on the rightmost stem. To analyze these changes as resulting from a floating tone, we would need a rather stipulative rule that causes the tone to spread leftward in the case of compounds and rightward in the case of genitives. Further, there is no independent evidence for floating tones in the language, so I argue that a floating tone analysis of compounds unnecessarily complicates the system, which is better accounted for with tonal overlays.

<sup>4</sup> Final [u] in Tommo So can be analyzed as being epenthetic, and especially on {H} stems, they are surface underspecified for tone; on {LH} stems, the H can shift onto the [u]. The exact f0 contour on the epenthetic vowel depends on the tonal context, and hence I leave the tone unmarked. See McPherson (2011) for further discussion.

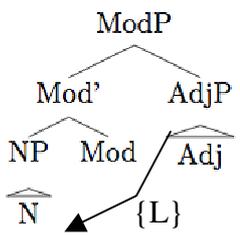
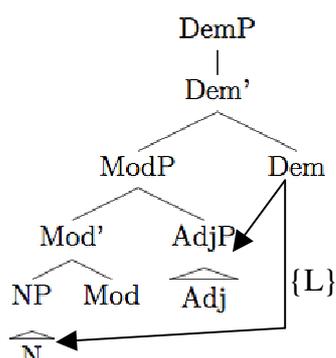
Interestingly, this same pattern of tone lowering on non-final stems is found in N+Adj constructions (3a), including recursive lowering when more than one adjective is present (3b):

- (3) a. **màlbá** + **pílu** → **màlbà<sup>L</sup>** **pílu**  
 gun white gun white  
 ‘white gun’
- b. **màlbà<sup>L</sup>** **pílu** + **kàndá** → **màlbà<sup>L</sup>** **pílù<sup>L</sup>** **kàndá**  
 gun. white new gun white new  
 ‘new white gun’

At first glimpse, this parallel tonal behavior raises the question of whether the underlying structure of N+Adj constructions and CCs is the same; after all, both involve a modified noun. However, the compounds in (2) are right-headed, meaning that it is the head noun that retains its lexical tone, whereas in (3), the head noun is on the left and the modifier retains its lexical tone.

A similar tonal situation is found in the distantly related language Akan (Marfo 2005). In this language, both N+N compounds and some N+Adj constructions become what Marfo calls “phrasal words”, in which N<sub>1</sub> takes {L} tone. He analyzes this as a rule of H-tone deletion operative on any non-final H tones in the phrasal word. As I will show below, a similar rule holds in Tommo So compounds, but I argue that the surface similarity with N+Adj constructions is just a surface coincidence.

I have argued elsewhere (McPherson 2010, McPherson, in prep) that tone lowering in N+Adj constructions (as well as N+Dem and N+Relative clause) is a consequence of the syntactic relation c-command, whereby a syntactically higher modifier is able to reach down and control the tone of the noun (and any other modifiers) it modifies. Consider the following:

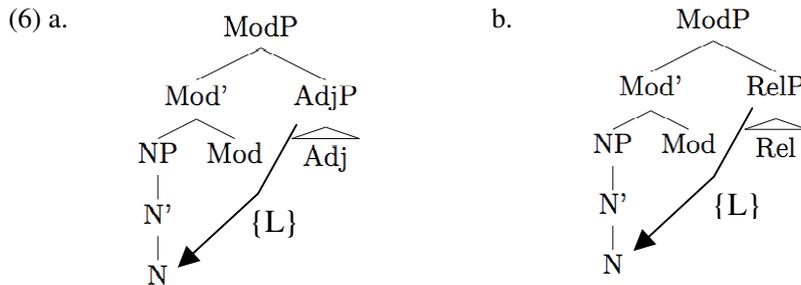
- (4) a.  **jàndùlù<sup>L</sup>** **gém** (cf. **jàndúlu**)  
 donkey black  
 ‘black donkey’
- b.  **jàndùlù<sup>L</sup>** **gém<sup>L</sup>** **nǒ**  
 donkey black this  
 ‘this black donkey’

In (4a), we see that an AdjP asymmetrically c-commands the head noun N, and in this configuration, it is able to impose a {L} tonal overlay. In (4b), we find that when we add a higher modifier, in this case a demonstrative, that higher modifier c-commands both the adjective and the noun, resulting in the tone lowering of both.<sup>5</sup>

In the cases of N+Adj and N+Dem, where the modifier is phrase final, a purely phonological analysis with a constraint against non-final H tones in modified phrases could account for the data. However, the same cannot be said for N+Relative constructions. This is due to the fact that Tommo So has internally headed relative clauses (with optional head fronting) and still tone lowering picks out this head in the middle of a phrase. This is shown in (5), where **nìŋé** ‘sauce’ has its /LH/ tone overwritten with {L} when acting as the head of a relative clause (delineated by square brackets). The sentence in (5b) shows that **nìŋé** can be optionally fronted but still takes a {L} overlay:

- (5) a. [Yàà-ná = ge                      nìŋé<sup>L</sup>      mí = ð                      jýè-m-ì = ge]  
 woman-HUMSg=DEF      sauce      1sgPRO=OBJ      eat-CAUS-PFV.REL=DEF  
**àmám = be.**  
 bitter=be.PFV  
 ‘The sauce the woman made me eat was bitter.’
- b. [Nìŋé<sup>L</sup> yàà-ná = ge                      mí = ð                      jýè-m-ì = ge]  
 sauce woman-HUMSg=DEF      1sgPRO=OBJ      eat-CAUS-PFV.REL=DEF  
**àmám = be.**  
 bitter=be.PFV  
 ‘The sauce the woman made me eat was bitter.’

In (5a), the object head of the relative clause, **nìŋé** ‘sauce’, is left in situ, surrounded by constituents containing H tones in both the relative subject and the relative participle; still, it undergoes tone lowering, showing that this tonal modification cannot be motivated by a need to delete all but the final H. In (5b), the object head is moved to the front of the relative clause and the tone lowering facts remain the same. The following clause still retains many H tones, and so a constraint against non-final H tones cannot be active. The tone lowering facts with relative clauses and other modifiers can be unified under the c-command analysis, if we assume that RelP and AdjP are both specifiers of ModP, shown in (6).

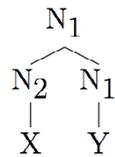


<sup>5</sup> The distinction between the phrase level c-command in (4a) and the head level c-command in (4b) can be empirically motivated by the fact that all adjectives impose tone lowering, while only certain determiners do, suggesting that tone lowering is a property of the head in the latter case and a property of the structural position in the former.

In (6a), we see an AdjP situated in the specifier of a ModP (modifier phrase). From this position, it c-commands the lower noun and is able to impose a {L} overlay. A parallel structure is shown in (6b), where a RelP is the modifier hosted by the ModP. Since it c-commands the lower noun, it also imposes a {L} overlay. I leave the exact syntax of internally headed relative clauses in Dogon, including how the head gets spelled out inside of the relative clause, to later work. The conclusion to draw from these trees is that the tone lowering seen on a noun when it is modified is a consequence of the underlying structural configuration and not a phonological constraint on the surface sequence of tones, since in the case of internally-headed relative clauses, this constraint predicts the wrong results.

While a structural account based on c-command can account for the tonal patterns in phrasal constructions with modifiers, I argue that it cannot explain the tone lowering found in compounds. That is, even though the surface result is the same (tone lowering non-final stems) in both compound nouns and N+Adj constructions, the motivation for the {L} overlays differs. I argue that compound tone lowering is in fact phonologically motivated, subject to a constraint like that found in Akan that mandates that non-final H tones are replaced with {L}. This constraint applies because the compound noun is contained under a single N node:

(7) *Branching structure of CCs*



The tree in (7) illustrates an N node (the syntactic head of the compound's NP) branching into two further N nodes representing the compound's constituent stems. The semantic head of the compound is  $N_1$ , on the right, whose features percolate up to the dominating syntactic head, coindexed with  $N_1$ . If the compound were left-headed, then the syntactic head node would be coindexed with  $N_2$ .

The consequence of this structure is that both compound nouns and simplex nouns are represented in the syntax by an N head, the difference being that compound nouns have another layer of structure below this head. I argue that the tonotactic restrictions on Tommo So nouns are operative on the head level, meaning that compounds and simplex nouns are subject to the same well-formedness constraints on tone. In section 1, I showed that the only licit tonal melodies for native nouns are /H/ and /LH/, and so when two stems, each carrying their own underlying /H/ or /LH/, come together to form a compound, the only way to guarantee that the output of the compound has a licit tonal structure is to replace the non-final stems' lexical tone with {L}. This is illustrated in the following table:

(8) {L} on  $N_1$  ensures overall licit tone pattern

$N_1$	$N_2$	/LH/	/H/
{L}		✓L-LH	✓L-H
/H/		*H-LH	✓H-H
/LH/		*LH-LH	✓LH-H

The potential tonal patterns of non-final stems ( $N_1$  in a two-stem compound) are represented in bold along the side; the tonal patterns of the final stem ( $N_2$  in a two-stem compound) are shown in italics along the top. While allowing  $N_1$  to retain its lexical tone results in tonally well-formed words if  $N_2$  is of the /H/ class (rightmost column), this creates ill-formed words when  $N_2$  is /LH/, shown by the starred forms H-LH and LH-LH in the second column. Enforcing either an all {H} overlay or a {LH} overlay on the first noun still runs into trouble, creating either illicit H-LH or LH-LH melodies. Only a {L} overlay placed on non-final stems ensures that the concatenation of any and all stems abides by word-level tonotactics.<sup>6</sup> This shows that the tone lowering seen in Tommo So CCs is not a parallel of that seen in other modified NPs but is instead motivated by a phonological well-formedness constraint on words, represented in the syntax by an N head.

## 2.2 HEADEDNESS

A well-known rule in the study of compounds is Williams's (1981) Right-hand Head Rule, which states that it is a universal principle that the semantic head of a compound be the rightmost stem. It is this right branch that percolates its features up to the dominating N node, as shown in (7). While the majority of CCs in Tommo So are right-headed, the existence of left-headed compounds in the language calls the universality of the rule into question, as others, such as Selkirk (1982), have done in the past. In the following sub-sections, I present all head configurations, showing that while semantics does not give us complete predictability, certain kinds of compounds do have a higher probability of being left-headed.

### 2.2.1 Right-headed CCs

Most CCs are right-headed, in conformity with the Right-hand Head Rule. Some illustrative examples are listed in (9):

(9)	Compound		Literal translation	Gloss
	<b>sùkòrò<sup>L</sup> jím</b> (cf. <i>súkòrò</i> )		sugar <sup>L</sup> disease	'diabetes'
	<b>gùnjù<sup>L</sup> sèlé</b> (cf. <i>gùnjú</i> )		toad <sup>L</sup> chain	'chain of eggs left by a toad'
	<b>àná<sup>L</sup> díí</b> (cf. <i>àná</i> )		rain <sup>L</sup> water	'rainwater'
	<b>kùrà<sup>nL</sup> sùmì-y</b> (cf. <i>kùrà<sup>n</sup></i> )		electricity <sup>L</sup> rope-DIM	'electricity wire'

In all cases, it is the final stem (the one with lexical tone) that heads the compound. Thus, 'diabetes' is a kind of disease, a 'chain of eggs left by a toad' is a kind of chain, 'rainwater' is a kind of water, and 'electricity wire' is a kind of little rope, or wire.

<sup>6</sup> We may wonder why this {L} overlay is not only applied in those cases where the combination of stems would result in an illicit tonal melody. That is, why change *àná díí* 'rainwater' into **àná<sup>L</sup> díí**, when both have an overall LH melody on the surface? The answer may lie in a constraint on the uniform realization of non-final stems. That is, if we only repaired illicit resulting tonal melodies, then in the case of *àná díí* 'rainwater', *àná* 'rain' would retain its lexical /LH/, but in **àná<sup>L</sup> gám-iyé** 'thunder' (literally 'rain rumbling'), it would need to take {L} to avoid the creation of a LH-LH melody. This results in a lack of uniformity between the realization of *àná* 'rain' in non-final position in compounds. It may be easier for learners to learn a rule that reduces *all* non-final stems to {L}, regardless of the tone of the final stem.

We saw above in ‘leather covering on the butt end of a rifle’ that the right-hand stem, the head, can be modified by another compound. In rare cases, it can also be modified by a N+Adj phrase, as in:

- (10) **ànà-m<sup>L</sup>**            **pè-m<sup>L</sup>**            **tànnàá** (cf. **ànà-m<sup>L</sup>** **pě-m** ‘old men’)  
 man-HUMPL    old-HUMPL staff  
 ‘kind of staff used by old men’

Here, an ‘old man staff’ does refer to a kind of staff (the right-hand head), characterized by being used by old men. We see that in the modifier, ‘old men’, both the noun and the adjective are inflected for the human plural using the suffix **-m**.

Traditional wisdom on compounds states that there should be no internal inflection, since this would be in violation of ordering restrictions (compounding as a derivational process should precede inflectional processes) (Allen 1978). However, speakers’ judgments point to the fact that irregular inflection inside of compounds is much more acceptable than regular inflection. Thus, “mice-catcher” is judged to be a better formation than “rats-catcher”. The Tommo So data uphold these facts. The human inflectional suffixes are in fact highly restricted, occurring on only a small set of nouns and an even smaller set of adjectives. While we find examples like (10) with this irregular inflection, there are no CCs that contain the regular plural clitic, **mbe**. This is in contrast to pseudo-genitive compounds, as we will see in section 4.

### 2.2.2 Left-headed compounds

Left-headed compounds, which violate the Right-hand Head Rule, tend to fall into two categories. The first are what I call “purposive compounds”, where in a compound form X Y, the translation would be “X for Y-ing”; that is, the non-head designates the purpose of the left-hand head. However, as the Y in these compounds is a deverbal noun, this places it in the synthetic compound category, and I will not discuss them here. For further description, see McPherson (forthcoming). The second category of left-headed compounds is what I call “elaborative compounds”, where the non-head elaborates on the left-hand head. For example:

- (11) a. **gòrò<sup>L</sup>**    **bèèré**            (cf. **góró** ‘hat’)  
 hat    beret  
 ‘beret’
- b. **pònnù<sup>L</sup>**    **àndéngé** (cf. **pónnu** ‘pants’)  
 pants    pant.kind  
 ‘short baggy pants’
- c. **bòmbòm<sup>L</sup>**    **yògú**            (cf. **bòmbóm** ‘candy’)  
 candy    tree.sp  
 ‘chewing gum’<sup>7</sup>

In (11a), ‘beret’, the left-hand stem simply means ‘hat’, while the right-hand stem, French loan **bèèré**, elaborates on this concept. A similar situation is found in (11b), where ‘short baggy pants’ are a kind of pants, and once again ‘pants’ is the left-hand stem. **Àndéngé** specifies the particular kind of pants in question. (11c) is an example where the right-hand stem really is more of a modifier, that is, the compound refers to

<sup>7</sup> This sticky candy is referred to with the name of a tree species, **yògú**, known for its sticky sap.

a kind of candy (the left-hand stem) characterized by having the sticky quality of the sap of the **yògú** tree (the right-hand stem). That is, **yògú** on its own does not refer to a type of candy, in the way that **bèéré** and **àndéngé** refer to articles of clothing.

As these examples suggest, left-headed elaborative compounds are particularly common in lexical realms like clothing and food. Nonetheless, plenty of compounds in these categories are right-headed, such as **jùlòm<sup>L</sup> góró** ‘wide straw-brimmed hat’, where ‘hat’ is the right-hand stem. While the semantic category can make it more likely that a compound would be left-headed, it cannot guarantee it.

### 2.2.3 Externally-headed compounds

The preceding two sections have addressed compounds whose heads are internal, whether the heads are on the right or on the left. The final class of CCs are those externally-headed or exocentric compounds, whose heads are neither of the overt stems but rather a third null element understood to be modified by the overt compound. Being null, this external head cannot carry a wide variety of meanings or it would quickly become very hard for a speaker to interpret. All attested cases of exocentric CCs are of the same form: The overt stems are a compound indicating a person, with the external head meaning roughly ‘state of being (person)’. For example:

- (12) a. **gùlònnò<sup>L</sup> náá** (cf. **gúlónnó** ‘post-partum quarantine’)  
 quarantine mother  
 ‘state of being in post-partum quarantine’
- b. **hòðlál<sup>L</sup> bàṅáá** (cf. **hóðlál** ‘trust’)  
 trust owner  
 ‘trustworthiness (state of being a trustworthy person)’

We see that the overt stems refer to people, be they a ‘quarantined mother’ or an ‘owner of trust’. Nonetheless, these compounds are not interpreted as referring to these people; they rather refer to an abstract state of being such a person. As we will see in section 3, all it takes is a tone change to remove the external head and identify the rightmost stem as the human head of the compound.

## 2.3 SUMMARY OF CCS

This section has described the tonal structure and headedness of Tommo So canonical compounds. I have argued that the tone lowering operative on non-final stems in a compound follows from the fact that CCs have a branching head structure. This means that the whole compound is contained in a single N node, just as a simplex noun is, and as such, both are subject to tonotactic constraints requiring that words take either a /H/ or /LH/ tonal melody. This is in contrast to N+Adj and other modifier phrases, where tone lowering is motivated by a structural configuration in which the modifier c-commands the noun; tone lowering in this case cannot be explained by a purely phonological constraint.

The distribution of head types in Tommo So canonical compounds is also not surprising, given the structure. Most compounds are right-headed, which conforms to Williams’s (1981) Right-hand Head Rule, which, though not universal, is certainly a strong tendency cross-linguistically. However, left-headed and exocentric compounds are also attested. Given the fact that headedness is simply determined by percolating the features of the head N up to the dominating N node, there is no foolproof reason

why this must be the right branch. This is in contrast to pseudo-genitive compounds below, where the structure itself correctly predicts that only right-headed compounds will fall into this class.

### 3. PSEUDO-GENITIVE COMPOUNDS

Pseudo-genitive compounds (henceforth PCs) derive their name from their tonal pattern, which is identical to that of a genitive construction. They are only right-headed and are laxer in their allowance of inflection inside of the compound.

#### 3.1 TONAL SHAPE

If CCs can be characterized as having the tone pattern  $X^L Y$ , then PCs can be characterized as the opposite:  $X Y^L$ . That is, in a simple two stem compound, the initial stem retains its lexical tone and the final stem takes {L}. For example:

- (13) a. **bílím kèè<sup>n</sup>-kèè<sup>nL</sup>** (cf. **kèè<sup>n</sup>-kéé<sup>n</sup>** ‘bug’)  
 manure Red-bug  
 ‘dung beetle’
- b. **íné bònnàà-ý<sup>L</sup>** (cf. **bònnàà-ý** ‘little whip’)  
 iron whip-DIM  
 ‘metal wire’
- c. **nàá=mbe pèndè<sup>L</sup>** (unknown citation tone)  
 cow=PL anthrax  
 ‘bovine anthrax’

As we can see, these compounds form phrasal tone patterns, with sequences like H-L and LH-L, unlike in the single word tonology of the CCs. The tone pattern is in fact identical to that of true genitives. Compare the compounds in (13) to the genitive constructions in (14):

- (14)a. **íí=ge kèè<sup>n</sup>-kèè<sup>nL</sup>** (cf. **kèè<sup>n</sup>-kéé<sup>n</sup>** ‘bug’)  
 child=DEF Red-bug  
 ‘the child’s bug’
- b. **Sáná bònnàà<sup>L</sup>** (cf. **bònnàá** ‘whip’)  
 Sana whip  
 ‘Sana’s whip’
- c. **yàá-m=ge=mbe jìm<sup>L</sup>=ge=mbe** (cf. **jím** ‘disease’)  
 woman-HUMPL=DEF=PL disease=DEF=PL  
 ‘the women’s diseases’

As we will see in the tree in (18), possessive tone is like modifier tone in that it is structurally motivated, with the possessor c-commanding the possessed noun. One difference between PCs and true genitives is that the “possessor” (the initial stem) in a PC cannot take the definite marker **ge**, which is expected, since the modifying element in a compound is meant to be non-referential (i.e. the ‘carrot’ in ‘carrot cake’ does not refer to a particular carrot). Nonetheless, we do see in (13) above that the regular plural clitic **mbe** can be used in these compounds, further corroborating their status as phrasal.

The examples in (13) are all simple PCs with two stems. The possibilities for complex PCs are more numerous than those for CCs (where recursive tone lowering of non-final stems is the only option). The possessor may itself be a CC (15a); the possessed portion (the head) may be a CC (15b); the possessor may be a PC (15c); the possessor may carry a modifier (15d); or most strikingly, the possessor may be a fully inflected phrase (15e):

- (15) a. [[CC] N<sup>L</sup>]  
 [[**ànù<sup>L</sup>** **pàgá**] **dòndò<sup>L</sup>**] (cf. **dóndó** ‘long bone’)  
 leg thigh long.bone  
 ‘thighbone’
- b. [N [CC<sup>L</sup>]]  
 [**ámhá** [**sìràà<sup>L</sup>** **bàrà-y<sup>L</sup>**]] (cf. **sìràà bàrà-ý** ‘snuff box’)  
 God snuff box-DIM  
 ‘kneecap (literally ‘God’s snuff box’)
- c. [[PC] N<sup>L</sup>]  
 [[**sáǵá** **dèmbèlè<sup>L</sup>**] **kàǵnàà<sup>L</sup>**] (cf. **kàǵnàà** ‘granary’)  
 space.under.granary platform granary  
 ‘granary built on a raised platform’
- d. [[N Mod] N<sup>L</sup>]  
 [[**ndé** **kém**] **ndè<sup>L</sup>**] (cf. **ndé** ‘person’)  
 person all person  
 ‘most beloved person’ (literally ‘everybody’s person’)
- e. [[S] N<sup>L</sup>]  
 [[**náá** **sê-m**] **pònnù<sup>L</sup>**] (cf. **pónnu** ‘pants’)  
 mother have-1SG.S pants  
 ‘type of pants’ (literally ‘pants of “I have a mother”’)

The different combinations of possessor and possessed give us a wide range of phrasal tone patterns, but always ending in the {L} overlay characteristic of possession. Example (15e) is most curious, where a phrase fully inflected with 1sg subject agreement acts as the possessor of **pónnu** ‘pants’. This, along with examples like (13c) containing the regular plural marker, suggests that PCs in Tommo So are lexicalized syntactic phrases, as Di Sciullo & Williams (1987) have argued for many French compounds like **cessez-le-feu** ‘ceasefire’.

If we assume that these PCs are indeed lexicalized syntactic phrases, then they can have the exact same structure as true possessive phrases. The result is that the tonal overlap with genitive constructions is in this case non-coincidental, deriving from the same syntactic configuration. Nonetheless, we face a complication here in that Tommo So has a split between alienable and inalienable possession, which (as I have argued in McPherson 2010 and McPherson in prep) derive from different structures. Which possessive structure should the PC take?

Simply the tone of the possessed noun cannot distinguish between the constructions, since a non-pronominal possessor (the kind of possessor seen in PCs) in both constructions assigns a {L} overlay to the possessed noun. The difference in Tommo So is realized when an adjective modifies the possessed noun. If the

possession is alienable, the adjective will also receive a {L} overlay; if inalienable, it will retain its lexical tone:

- (16) a. **Sána** **jàndùlù<sup>L</sup>** **kòmmò<sup>L</sup>**  
 Sana donkey skinny  
 ‘Sana’s skinny donkey’
- b. **Sána** **bàbè<sup>L</sup>** **kómmó**  
 Sana uncle skinny  
 ‘Sana’s skinny uncle’

In (16a), **kómmó** ‘skinny’ is overwritten with {L} because **jàndùlu** ‘donkey’ is an alienably possessed noun. In (16b), it retains its tone, since **bàbé** ‘uncle’ is inalienably possessed.<sup>8</sup> If we put an adjective after a PC, we would get a clear indication of the underlying structure from the surface tone.

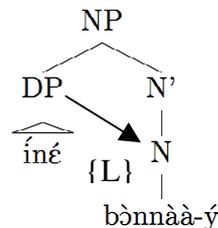
When a PC is modified, we find that the tone patterns mirror those of inalienable possessives, like that in (16b); both the possessor and the adjective retain their lexical tones, while the possessed noun (head of the compound) undergoes tone lowering. For example:

- (17) a. **íné** **bònnàà-ý<sup>L</sup>** **kàndá**  
 iron whip-DIM new  
 ‘new wire’
- b. **bílím** **kèè<sup>n</sup>-kèè<sup>nL</sup>** **bánu**  
 manure Red-bug red  
 ‘red dung beetle’

In (17a), we see that both the pseudo-genitive possessor **íné** ‘iron’ and the adjective **kàndá** ‘new’ retain their lexical tone; similarly in (17b), only the head noun, **kèè<sup>n</sup>-kèè<sup>n</sup>** ‘bug’ undergoes tone lowering. These tonal data point to PCs taking the structure of inalienable rather than alienable possessives.

Assuming this structure for Tommo So PCs, we can model the PC **íné bònnàà-ý<sup>L</sup>** ‘metal wire’, literally ‘iron’s little whip’, as follows:

- (18) Inalienably possessed structure of ‘metal wire’



The possessor **íné** ‘iron’ is in its own DP in the specifier of **bònnàà-ý** ‘little whip’, the head noun’s, NP. It c-commands the N and imposes a {L} overlay, just as we saw with adjectives and relative clauses in section 2. Given this structure, the whole compound is contained within a single NP, capturing the close relationship between the stems of a PC. Alienable possession is often considered to have a larger structure

<sup>8</sup> In the Dogon languages, the class of inalienably possessed nouns is rather small, comprised of only kinship terms and a few other relations like ‘namesake’. Body parts and even wives and children are alienably possessed.

(Dobler 2008), with the possessor being in a higher phrase such as PossP, that would disrupt this close relationship.

What is interesting here is that PCs are using a structure available in the syntax for inalienable possession even when the stems involved would not typically be possessed this way. That is, we do not see an exact lexicalization of a possessive phrase, or else ‘metal wire’ would end up in a larger structure containing PossP. Instead, these stems borrow the more condensed structure of inalienable possession, which draws them together into a compound while still maintaining a phrasal structure (as opposed to the head-level CCs) and the same surface tonal pattern (X Y<sup>L</sup>).

### 3.2 HEADEDNESS

Another difference between CCs and PCs is the position of the head. While CCs are very diverse in terms of headedness, PCs are exceptionlessly right-headed. In fact, the very same exocentric compounds that we saw in (12) switch to having a right-hand head when the tone pattern switches from X<sup>L</sup> Y to X Y<sup>L</sup> (i.e., from CC to PC):

- (19) a. **gùlònnò<sup>L</sup>**      **náá**  
 quarantine      mother  
 ‘state of being in post-partum quarantine’
- gúlónnó**      **nàà<sup>L</sup>**  
 quarantine      mother  
 ‘a woman who is in post-partum quarantine’
- b. **hòòlál<sup>L</sup>** **bàṅáá**  
 trust      owner  
 ‘trustworthiness’
- hóòlál**      **bàṅàà<sup>L</sup>**  
 trust      owner  
 ‘trustworthy person’

The initial member of each pair above represents the exocentric CC compound, referring to the state of being the person indicated by the overt stems in the compound. In the second member, the tone switches to that of a PC, and accordingly, the compound becomes right-headed. That is, in (19a), **gúlónnó nàà<sup>L</sup>** with PC tone refers to the ‘quarantine mother’, a woman who is in post-partum quarantine. In (19b), **hóòlál bàṅàà<sup>L</sup>** refers to the ‘owner of trustworthiness’, that is, a trustworthy person.

Looking at the structure of PCs proposed in (18), the exclusivity of right-hand heads falls out naturally from the fact that the right-hand stem is syntactically the head of the NP that contains both it and the modifying “possessor”. This is in contrast to CCs, where the head node N simply branches into two further N nodes, with no a priori syntactic reason for the right or the left to percolate its features to the dominating N node.<sup>9</sup>

The question still remains: Why do we not see any exocentric PCs? Presumably there is no syntactic reason why the head N of the possessive NP cannot be null, if a null head is possible in CCs. One possible reason behind this gap is a functional one

<sup>9</sup> Clearly the preponderance for right-headed compounds cross-linguistically suggests a universal bias towards percolating the features of the right branch up, but this asymmetry is not coded in the syntax to the extent that the N of a possessive construction is head of its NP.

based on the confusability of the tonal output. Recall that the tone of possessive constructions is  $X Y^L$ , where X is the possessor and Y the head of the possessive NP. If the possessed noun were null, there would be no difference between a headless possessive construction and simply the possessor as its own NP:

- (20)  $y\grave{a}\grave{a}-n\acute{a}$   $\emptyset^L$  =  $y\grave{a}\grave{a}-n\acute{a}$   
 ‘woman’s thing’ ‘woman’

A possessor ‘woman’ without a L-toned possessed noun following it sounds just like ‘woman’ on its own. Of course, even the headless CCs we see never consist of a single overt stem, since in that case, too, the intended compound would be indistinguishable from a simple noun. So why is it that the PC  $h\acute{o}\grave{d}l\grave{a}l\ b\grave{a}\grave{n}\grave{a}\grave{a}^L$  ‘trustworthy person’ cannot be interpreted as exocentric ‘state of being a trustworthy person’? Another possible explanation is that headless possessive phrases in Tommo So cannot be bare. That is, they always contain a possessive clitic **mo** that signals to the listener that the utterance is meant to be possessive. For example, the attempted headless ‘woman’s thing’ from (20) would be pronounced  $y\grave{a}\grave{a}-n\acute{a} = mo$ . In contrast, headless modifiers (adjectives, demonstratives, relative clauses) can stand alone. Thus, **b\acute{a}nu** ‘red’ could stand alone meaning ‘red thing’. Given this situation, the combination of stems in the PC  $h\acute{o}\grave{d}l\grave{a}l\ b\grave{a}\grave{n}\grave{a}\grave{a}^L$  ‘trustworthy person’ cannot be interpreted as a headless possessive construction since there is no clitic **mo** signaling that this is the case. Why it is that we never see PCs with **mo** is less clear.

### 3.3 INTERIM SUMMARY

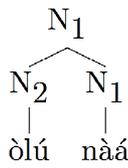
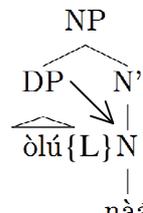
Sections 2 and 3 have investigated CCs and PCs in Tommo So, showing that the tone patterns of each class suggest an underlying structure of a branching head for CCs and an inalienable possessive NP for PCs. That CCs are a single N node while PCs are phrasal is further upheld by the fact that phrasal PCs allow regular inflection inside of the compound and predetermine the head position.

## 4. VARIABILITY

A final interesting fact about CCs and PCs is that with some combinations of stems, speakers will vary as to which tonal pattern they apply. On one occasion, a speaker may pronounce the compound with PC tone ( $X Y^L$ ) and on the next with CC tone ( $X^L Y$ ). An example of this variation is shown in (21):

- (21)  $\grave{o}l\acute{u}$   $n\grave{a}\grave{a}^L$  (PC) ~  $\grave{o}l\grave{u}^L$   $n\acute{a}\acute{a}$  (CC)  
 bush cow bush cow  
 ‘water buffalo’

There is no change in meaning associated with the change in tone. I argue that the differences in tone pattern arise from variation in the structural position of the modifying noun; speakers can head-adjoin it, forming a CC, or they can put it in the specifier of the NP, forming a PC. These two constructions are illustrated by the trees below:

- (22)a.  b. 

The pattern of tone lowering in (22a) would be assigned based on word-level tonotactics, whereas the {L} overlay in (22b) is assigned based on the structural configuration of the possessor and the possessed noun. The result is that the CC in (22a) surfaces as **òlù<sup>L</sup> nàá** and the PC in (22b) as **òlú nàà<sup>L</sup>**.

It is not clear how pervasive this variation is, but it seems particularly common with non-final **òlú** ‘field’, suggesting that the variation may be a property associated with non-heads rather than heads.

## 5. CONCLUSION

This paper has investigated the two tonal classes of root compounds found in Tommo So, showing that tone can be used as a window onto underlying structure. In the case of CCs, the non-final stems in the compound are subjected to tone lowering in order to satisfy a tonotactic constraint on nouns; this constraint is able to apply to CCs because they are structurally a branching N node, meaning they are contained in the same level of structure as simplex nouns. While it is usually the features of the right-hand branch that percolate up to the dominating N node, CCs also allow the features of the left branch to percolate up or even the features of a null head.

PCs, on the other hand, have a phrasal tone pattern identical to that found in possessive constructions in the language. When an adjective is added to the PC, tonal evidence points to the fact that these compounds are treated as inalienable rather than alienable constructions, possibly because then the whole compound is contained under a single NP node. The fact that these compounds are lexicalized phrases is further clear from the fact that regular inflection can occur inside of the compound. The {L} tone overlay is imposed by a possessor that c-commands the noun, the same mechanism employed by modifiers like adjectives and relative clauses to impose {L} overlays on the nouns they modify.

The tonal overlays found in Tommo So compounds are characteristic of the tonal morphology of the Dogon language family as a whole. While many West African languages make wide use of floating tones and other sorts of concatenating tonal morphemes, the Dogon languages typically use tonal overlays in both the NP and in verb inflection that completely erase the stem’s lexical tone. These tonal overlays are intricately related to underlying structure in sometimes surprising ways (Heath & McPherson, forthcoming). This paper has shown one such case, where the underlying syntactic structure of a compound has direct consequences for its tonal realization. Tonal overlays resulting from phonological constraints are attested in other languages of West Africa, such as the phrasal word rule of Akan (Marfo 2005) or *compacité tonale* in Bambara (Creissels 1988, Green 2010, among others), but it remains to be seen whether structurally-motivated overlays of the sort seen here are unique to the Dogon languages or are more widespread among the languages of the region.

## ABBREVIATIONS

CAUS	Causative	PC	Pseudo-genitive compound
CC	Canonical compound	PFV	Perfective
DEF	Definite	PL	Plural
DIM	Diminutive	REL	Relative
HUM	Human	S	Subject
<sup>L</sup>	(tone lowering)	SG	Singular
OBJ	Object		

## REFERENCES

- Allen, Margaret. 1978. Morphological investigations. Ph.D. thesis, University of Connecticut.
- Cahill, Michael. 2007. More universals of tone. SIL Electronic Working Papers, 2007-007.
- Creissels, Denis. 1988. Le système tonale du bambara standard. Grenoble: Centre de Linguistique Africaine de l'Université Stendhal.
- Di Sciullo, Anna Maria and Edwin Williams. 1987. On the Definition of Word. Cambridge, Mass: MIT University Press.
- Dobler, Eva. 2008. The morpho-phonology of (in)alienable possession. Proceedings of 2008 annual conference of the Canadian Linguistics Association.
- Green, Christopher. 2010. Prosodic phonology in Bamana (Bambara): Syllable complexity, metrical structure, and tone. Doctoral dissertation, Indiana University.
- Heath, Jeff and McPherson, Laura. Forthcoming. Tonosyntax and reference restriction in Dogon NPs. University of Michigan and UCLA. Language.
- Marfo, Charles. 2005. Aspects of Akan grammar and the phonology-syntax interface. Ph.D. thesis, Hong Kong University.
- McPherson, Laura. 2010. Dogon alienable and inalienable possessives: Evidence of structure-sensitive tone in the DP. Paper presented at the 41<sup>st</sup> Annual Conference of African Linguistics, Toronto.
- \_\_\_\_\_. 2011. Tonal underspecification and interpolation in Tommo So. M.A. Thesis, UCLA.
- \_\_\_\_\_. Forthcoming . A Grammar of Tommo So. De Gruyter Mouton.  
In prep. Replacive grammatical tone in the Dogon languages: a factorial typology. Ph.D. thesis, UCLA.
- Nissim, Gabriel. 1977. Je parle bamileke. Douala: Collège Libermann.
- Plungian, Vladimir. 1995. Dogon. Lincom Europa.
- Schuh, Russell. 2009. Ngamo tones and clitics. Ms., UCLA.
- Selkirk, Elisabeth. 1982. The Syntax of Words. Cambridge: MIT Press.
- Williams, Edwin. 1981. On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* 12: 245-274.
- Williamson, Kay. 1986. The Igbo associative and specific constructions. *The Phonological Representation of Suprasegmentals*. Koen Bogers, Harry van der Hulst and Maarten Mous (Eds.) 195-208. Dordrecht: Foris.