Given the fact that so many languages in Africa and elsewhere in the world are poorly studied, there is obviously still a lot to be learned about the typology of languages. (Dimmendaal 2011: 298).

Head position in nominal compounds

A lesson from Africa
Far away, in the West African village of Galim, there is a language called Nizaa.

It exhibits a feature which appears to be unique in the published literature.

This has ramifications for the typology of compounding…

- (and perhaps also for the classification of compounds)
The head in compounding

- A prototypical compound consists of two or more lexical roots, one of which is the “head”
  - (Here I quietly ignore co-compounds, “exocentric” compounds, and the question of exactly what we mean by the head)

- The head generally determines the (grammatical and semantic) “profile” of the compound
  - e.g. black bird, snail mail, football, apron string, taxi driver, futbalový štadión, lesopark

- One of the questions posed by language typologists is
  - Is there a linguistic universal that governs the position of the head in compounds?
Proposal #1: Fixed position

- "Righthand Head Rule" (RHR)
  - “The head of a morphologically complex word is the right-hand member of that word” (Williams 1981)
  - Works for English and other Germanic languages (mostly)

- Falsified by evidence from Romance, Hebrew, Celtic, etc.
  - **French**: bateau mouche ‘river boat’ < ‘boat’ + ‘fly’
  - **Hebrew**: delet bayit ‘house door’ < ‘door’ + ‘house’
  - **Welsh**: drws tŷ ‘house door’ < ‘door’ + ‘house’

- All of these languages have left-headed compounds
Proposal #2: Parametric variation

- The “principles and parameters” approach
  - Each language is either left-headed or right-headed
    - “Stando alle nostre conoscenze attuali, non sembra che una lingua possa avere liberamente composti con testa a sinistra e composti con testa a destra” (Scalise 1994)

- Falsified by two kinds of evidence
  1) Vietnamese: LH native tradition vs. RH Chinese tradition
     - xe lũaN ‘train’ < xeN ‘vehicle’ + lũaN ‘fire’
     - tô-quốc ‘fatherland’ < 祖国 zǔguó ‘ancestor’ + ‘country’
  2) Mandarin: RH nominal compounds vs. LH verbal compounds
     - 食品 shípǐnN ‘food’ < shíV ‘eat’ + pǐnN ‘product’
     - 开刀 kāidāoV ‘operate’ < kāiV ‘open’ + dāoN ‘knife’
Proposal #3: Canonical position

- Discounting contact phenomena, it seems there is
  - “a canonical position for the head in each compound type in a given language” (Scalise & Fabregas 2010)

Where “compound type” means “output category type”

Enter Nizaa…

- Reported to have a mixture of left-headed and right-headed nominal compounds
  - i.e. compounds of the same output category type
- Not attributable to language contact
Nizaa: A Mambiloid language

Mambiloid: Fam, Kwanja, Mambila, Mbongno, Mvanip, Ndoro, Ndunda, Nizaa, Njerep, Somyev, Twendi, Vute, Wawa
Sample of N+N nominal compounds

<table>
<thead>
<tr>
<th>Nominal Compound</th>
<th>Analysis: Noun + Noun</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>sìì yim</td>
<td>‘dispensary’</td>
<td>&lt; ‘house’ + ‘medicament’</td>
</tr>
<tr>
<td>sìì jaŋgide</td>
<td>‘school’</td>
<td>&lt; ‘house’ + ‘reading’</td>
</tr>
<tr>
<td>sìì ɓʉʉ</td>
<td>‘roof’</td>
<td>&lt; ‘house’ + ‘head’</td>
</tr>
<tr>
<td>bag ɓʉʉ</td>
<td>‘hat’</td>
<td>&lt; ‘clothing’ + ‘head’</td>
</tr>
<tr>
<td>nim ɓʉʉ</td>
<td>‘source/spring’</td>
<td>&lt; ‘water’ + ‘head’</td>
</tr>
<tr>
<td>sìŋw nim</td>
<td>‘duck’</td>
<td>&lt; ‘bird’ + ‘water’</td>
</tr>
<tr>
<td>yír nim</td>
<td>‘tear’</td>
<td>&lt; ‘eye’ + ‘water’</td>
</tr>
</tbody>
</table>
Research questions

- Is it really the case that Nizao exhibits no particular preference for either left-headed or right-headed nominal compounds?
- If so, can a semantic analysis provide an explanation?
Data

- Field notes collected by Rolf Theil (genannt Endresen) in 1983–84
- 1300-page word list (including notes)
- No access to native speakers
Methodology

- **Identify nominal compounds**
  - 534 candidates

- **Identify constituents**
  - 314 transparent
  - 116 semi-transparent
  - 104 opaque

- **Classify by input type**
  - Predominance (63%) of noun + noun compounds
  - 203 NN, 35 NA, 33 NV, ...

- **Determine head position**
  - 60-40 split between LH and RH compounds

- **Statistical analysis of recurrence patterns**
  - Heads recur significantly more often in left-headed compounds
  - Non-heads recur significantly more often in right-headed compounds

- **Semantic analysis...**
Analysing semantic relations

- Jackendoff (2010) not directly usable
  - Some functions too general, e.g. CLASSIFY(x, α)
  - Some functions too culture-specific, e.g. PROTECT(α, x)
  - Some functions too language-specific, e.g. HAVE(α, x)

- Warren (1978) and Ryder (1994) equally difficult to apply
  - (Adopting someone else’s classification is not easy)

- Devised own ad-hoc classification inspired by the above
  - Incorporated insights from knowledge modelling
  - The methodology is prescientific, but still valid

- The results were quite startling…
### Examples of semantic relations

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
<th>Component 1</th>
<th>Component 2</th>
<th>Relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ḋów nyíŋ</td>
<td>‘canine’</td>
<td>&lt; ‘dog’ + ‘tooth’</td>
<td></td>
<td>POSSESSION</td>
</tr>
<tr>
<td>sìì jaŋgide</td>
<td>‘school’</td>
<td>&lt; ‘house’ + ‘reading’</td>
<td></td>
<td>ACTIVITY</td>
</tr>
<tr>
<td>sìì ɓuụ</td>
<td>‘roof’</td>
<td>&lt; ‘house’ + ‘head’</td>
<td></td>
<td>PART</td>
</tr>
<tr>
<td>ḃag ɓuụ</td>
<td>‘hat’</td>
<td>&lt; ‘clothing’ + ‘head’</td>
<td></td>
<td>LOCATION</td>
</tr>
<tr>
<td>ḍó́ṅg niŋ</td>
<td>‘soup’</td>
<td>&lt; ‘sauce’ + ‘water’</td>
<td></td>
<td>APPEARANCE</td>
</tr>
<tr>
<td>sìŋw niŋ</td>
<td>‘duck’</td>
<td>&lt; ‘bird’ + ‘water’</td>
<td></td>
<td>LOCATION</td>
</tr>
<tr>
<td>yír niŋ</td>
<td>‘tear’</td>
<td>&lt; ‘eye’ + ‘water’</td>
<td></td>
<td>LOCATED</td>
</tr>
</tbody>
</table>
Distribution by semantic relation

<table>
<thead>
<tr>
<th>Relation type</th>
<th>Total</th>
<th>LH</th>
<th>RH</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOCATION</td>
<td>26</td>
<td>26</td>
<td>0</td>
</tr>
<tr>
<td>PURPOSE</td>
<td>13</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>ACTIVITY</td>
<td>9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>APPEARANCE</td>
<td>9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>IDEM</td>
<td>9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>STATE</td>
<td>7</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>GENDER</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>EVENT</td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>MATERIAL</td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>CONTENTS</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>PROPER</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>EMOTION</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>DOMAIN</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>FEATURE</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>FOOD</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PART</td>
<td>45</td>
<td>0</td>
<td>45</td>
</tr>
<tr>
<td>KIN</td>
<td>9</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>LOCATED</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>POSSESSION</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>PRODUCT</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>CONTAINER</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>RESPONSIBLE</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

- **22 relation types**
  - LOCATION, PURPOSE, ACTIVITY, etc.

- **Left-hand column**
  - 15 types, **all of which** occur in LH compounds only

- **Right-hand column**
  - 7 types, **all of which** occur in RH compounds only

- **One special case**
  - LOCATION and LOCATED are the inverse of each other
Summary of results

- **Right-headed compounds**
  - **PART-WHOLE** and **KINSHIP** relations predominate
    - also: **LOCATED**, **POSSESSION**, **PRODUCT**, **CONTAINER**

- **Left-headed compounds**
  - **APPEARANCE**, **GENDER**, **LOCATION**, **PURPOSE**, **ACTIVITY**
    - largely attributive

- Clearly, Nizao **does** exhibit both left- and right-headed compounds...

- And there is a **systematic difference** between the two, expressible in terms of semantic relations
Towards an explanation

- My explanation for this phenomenon is framed in terms of Langacker’s Cognitive Grammar
  - (Time does not permit me to go into the details)
  - (The paper is available at Academia.edu)

- It hinges on the notion of “mental access routes”
  - The patterns in Nizaa represent two such routes
  - They correspond (roughly) to the division of compounds into ”subordinate” and “attributive” (Scalise & Bisetto 2009)

- Essentially, in Nizaa
  - Subordinate compounds are right-headed
  - Attributive compounds are left-headed
A syntactically mixed language

- These two mental access routes are probably universal but they surface rather clearly in Nizɑɑ.
- Nizɑɑ is a little unusual (but far from unique) in that it has *postposed* modifiers and *preposed* possessors.
- From a diachronic perspective:
  - **LH compounds** develop from the one
    - lexicalization of noun phrases in which adjectival nouns act as postposed modifiers
  - **RH compounds** develop from the other
    - lexicalization of possessive phrases
In conclusion

- Nizɑɑ is very unusual (but surely not unique?) in having both left-headed and right-headed compounds in equal measure

- The two kinds are characterized by different semantic relations
  - “Subordinate” compounds are right-headed
  - “Attributive” compounds are left-headed

- We need to revise the Canonical Position Hypothesis
  - “Compound type” has to be defined in terms of semantic relation as well as output type
References


My PhD project (2015-2018)

**Nominal compounding in cross-linguistic perspective**

Please contact me if you are interested in the topic

<pepper.steve@gmail.com>

...especially if you already have (or intend to acquire) specialist knowledge about compounding in a lesser known language...