A Universal Typology of Truncation

Birgit Alber     Sabine Lappe
University of Verona    University of Siegen
birgit.alber@univr.it    lappe@anglistik.uni-siegen.de

1. What is truncation?
Derivation of a TRUNC morpheme from a base - mostly (but not necessarily) through shortening of the base:

(1)  a. English monosyllabic name truncation: size: \((H)\)
Dave - David
Fred - Frederick

b. Italian bisyllabic name truncation: size: \((σσ)\)
Fede - Federica
France - Francesca
Anto - Antonella

c. German i-Bildungen:
Andi - Andreas
Hansi - Hans

N.B.: prespecified final -i; the base is not necessarily shortened

1.1 GTT- deriving the template


The size of templates is not process-specific, but due to an 'Emergence of the Unmarked Ranking'.

(2) GTT-ranking for TRUNC templates
\(\text{MAX-IO} \gg \text{SRC} \gg \text{MAX-BT}\)

- MAX-IO: every segment in the input has a correspondent in the output
- MAX-BT: every segment in the base has a correspondent in TRUNC
- SRC: 'size restrictor constraints': prosodic markedness constraints (compatible with the constraint hierarchy of the language under investigation)

1.2 What is in the truncation - Anchoring

Most typically: the TRUNC morpheme preserves prominent material from the base:
- the first syllable of the base (s.1 a., b., d. above)
- the stressed syllable of the base:

(4) Italian bisyllabic name truncation - stress anchoring:
Césca - Francésca
Bérto - Robérto

- both the first and the stressed syllable of the base (cf.below)

e.g. Italian bisyllabic truncation: anchoring to the initial syllable

(5) SRC >> ANCHOR-\(σ_1\) >> ANCHOR-\(σ_{stress}\) >> MAX-BT

<table>
<thead>
<tr>
<th>/Federica/</th>
<th>SRC</th>
<th>ANCHOR-(σ_1)</th>
<th>ANCHOR-(σ_{stress})</th>
<th>MAX-BT</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (fe)</td>
<td>*!</td>
<td></td>
<td></td>
<td>derica</td>
</tr>
<tr>
<td>b. (fe.de)</td>
<td></td>
<td>*</td>
<td></td>
<td>rica</td>
</tr>
<tr>
<td>c. (ri.ca)</td>
<td></td>
<td></td>
<td>*</td>
<td>fede</td>
</tr>
<tr>
<td>d. (fe.de)(ri.ca)</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>
1.3 Our project - A universal typology of truncation

Given the assumptions of GTT and prominence-based anchoring:
- what systems does the factorial typology predict?
- what systems are there, out in the world?

Our data: truncation patterns extracted from the literature, tagged according to template type/size and anchoring

So far: 76 patterns, spread over 22 languages

Our claims:
- All patterns found in the data are predicted by the factorial typology (no undergeneration)
- Almost all patterns predicted by the factorial typology are attested in the data (some overgeneration, to be investigated)
- the factorial typology predicts atemplatic TRUNC morphemes which vary in size. This is an important difference between an OT-based account and previous, templatic approaches.

2. The factorial typology of Truncation

Three constraint-families involved in truncation:

Informally:
- SRC: generates the template (foot, syllable, etc.)
- ANCHOR-P: wants to preserve prominent material of the base
- MAX-BT: wants to preserve as much material as possible of the base

N.B., in order for truncation to happen at all we need for some SRC to dominate MAX-BT, hence the remaining logical possibilities are:

(6) the factorial typology

1. **SRCs >> ANCHOR-P >> MAX-BT**
   - Well-formed templates are created; if necessary, at the expense of anchoring

2. **ANCHOR-P >> SRCs >> MAX-BT**
   - Anchoring is satisfied; if necessary, at the expense of creating well-formed templates

3. **SRCs >> MAX-BT >> ANCHOR-P**
   - Preserving as many segments as possible from the base is more important than (good) anchoring

- What does the predicted interaction look like?
- Do we find evidence for this interaction?

**Pattern 1: SRCs >> ANCHOR**

... predicts that (good) anchoring may be sacrificed in order to ensure a well-formed template

(7) English monosyllabic nicknames (cf. 1c)
  - Hit *Me Mehitabel
  - Ron *A Aaron (non-rhotic)

(NB: anecdotal evidence, not a much-studied area so far)

**b. Pattern 2: ANCHOR >> SRCs**

... predicts that the well-formed template may be sacrificed in order to ensure good anchoring

(8) a. a radical case: Southern Italian Vocatives
  - Antoné *Anto *Nella Antonéla
  - Carmé *Cárme *Mela Carméla
  - Pa *Páo Páola

(to be discussed in section 3)

b. less radical: Spanish hypocoristics; template = syllabic trochee
  - Cánda *Can Cándida
  - Cénte *Cen Vicente
  - Géntfo *Xen Fulgencio

but:
  - Tíno Tín Agostín
  - Tiča Tíf Beatríz
  - čóna čon Encarnación

**Pattern 3: MAX-BT >> ANCHOR - overgeneration?**

Hypothetical case:
- a language with monosyllabic TRUNC-morpheme
- locus of anchoring shifts according to where a maximum of base segments can be preserved

First syllable anchoring: Car.mé.la --> Car (vs. *Me)

Stressed syllable anchoring: Ar.mán.da --> Man (vs. *Ar)
3. Zooming in on the SRCs

Size restrictor constraints in the literature: typically prosodic markedness constraints that determine the metrical structure of the language:

FT-BIN
PARSE-σ
ALL-FEET-LEFT/RIGHT
TROCH/IAMB
WSP

Work here is still in progress, but it is clear what should be done:
- the interaction of ALL of these (and some more) with ANCHOR-P and MAX-BT should be tested
- it should be checked whether the proposed ranking of the metrical constraints is compatible with the analysis of the metrical structure of the language

Some typologies we have uncovered:

3.1 Foot-templates

3.1.1 The maximal foot:

(9) PARSE-σ, ALL-FT-LEFT >> MAX-BT

generates the foot-template

wants maximal preservation

This is the 'normal' case: MAX-BT, if not inhibited, should always generate a maximal foot

(10) Central Alaskan Yupik proximal vocative (iambs, quantity-sensitive) (phonetic transcription follows Weeda 1992)

<table>
<thead>
<tr>
<th>Base</th>
<th>Pattern 1</th>
<th>Pattern 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apukasnaaq</td>
<td>Ag</td>
<td>Apuk</td>
</tr>
<tr>
<td>Nupiqaq</td>
<td>Nup</td>
<td>Nupix, Nupik</td>
</tr>
<tr>
<td>Aqivsnaq</td>
<td>Aqif</td>
<td></td>
</tr>
</tbody>
</table>

Foot-template (H) (L’H)

Pattern 2: generated by MAX-BT, foot should always be maximal
Pattern 1: cf. below

N.B.: absence of (LL) template predicted by activity of MAX-BT

More patterns exhibiting the maximal foot:

(11) Spanish hypocoristics – two patterns, one is maximal, one is not

Fernándo  Féman  Féma
Bernárdo  Bémar  Béma
Xesús  Xésus  Xésu

3.1.2 The wellformed foot

(12) Central Alaskan Yupik proximal vocative: first syllable heavy

apaqayaq  a.ìan
naqqałivia  ne.ŋaq
qakfasalria  qa.kaf

Foot-template (L’H) *H’H)

--> schwa-epenthesis can guarantee a wellformed (L’H) iamb and exclude a potential (H’H) foot

(13) WSP >> DEP, MAX-BT

excludes unstressed heavy syllables, as in (H’H)

More patterns exhibiting the wellformed foot:

(14) Japanese čan-suffixed hypocoristics

Midori  Mido-čan, Mii-čan
Yooko  Yoko-čan, Yoo-čan
Hanako  Hana-čan, Hač-čan, Hač-čan

(11) or (11)

3.1.3 More than one foot

(14) PARSE-σ >> MAX-BT >> ALL-FT-LEFT

all syllables must be parsed into feet more than one foot is possible, hence required by MAX-BT

i.e.: TRUNC may consist of more than one foot as long as no unparsed syllables remain
Japanese loan compound abbreviations (almost, maybe): consistently 4 moras:¹

(15) arubaito saron  aru-saro  'arbeit salon'
    (LL)(LL)
waado purosessaa  waa-puro  'word processor'
    (H)(LL)
rimoto koNtorooru  rimo-koN  'remote control'
    (LL)(H)

3.2 The monosyllabic template

Traditionally: the Minimal-Word-Template

Very common: s. Yupik, pattern 1, English nicknames (Fred-Frederick), Japanese Geisha-House Discretionary Names (o-Fuu-saN - Fukuda)

How can it be generated? Not a trivial question, given the activity of MAX-BT


(16) PROMMAX: every segment of the output is in a prominent position (first or stressed syllable)

(17) PROMMAX >> MAX-BT

  every segment must be in the first/stressed syllable = there is only one (maximally packed) syllable

(18) PROMMAX >> MAX-BT

<table>
<thead>
<tr>
<th>/Frederick/</th>
<th>PROMMAX</th>
<th>MAX-BT</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (Fred)</td>
<td></td>
<td>erick</td>
</tr>
<tr>
<td>b. (Fre.de)</td>
<td>de!</td>
<td>rick</td>
</tr>
<tr>
<td>c. (Fre.de)(rick)</td>
<td>de.rick!</td>
<td></td>
</tr>
</tbody>
</table>

¹ According to Ito (1990), Labrune (2002) loan compound abbreviations cannot be interpreted as separate truncation of each part of the compound since (H) abbreviations apparently do not exist in the realm of loan truncation.

3.2.1 The subminimal foot

The influence of syllable-structure constraints, e.g.

(20) NoCODA, PARSE-σ, ALL-FT-LEFT >> MAX-BT

  template is a foot ending in a syllable without coda

(9) Italian bisyllabic name truncation:

  Fe.de - Fe.de.ri.ca  *Fe.der
  Va.le - Va.len.ti.na  *Va.len

The effect of NoCODA in monosyllabic truncation:

(22) NoCODA, PROMMAX >> MAX-BT

  the template is a single syllable without coda, hence C(C)V

(23) Italian monosyllabic name truncation: (cf. also truncation in Zuñi)

  Cri - Cristina
  Fra - Francesca
  Lu - Luisa
  Ste - Stefania

Subminimal templates consisting of a degenerate foot have often been claimed not to exist, since they don't seem to correspond to the Minimal Word. However, they are predicted by GTT, at least if a language allows for degenerate feet in general, as Italian seems to do:

(24) (fa) 'do, 3 P Sg.'
    (e) 'be, 3 P. Sg.'
    (L)
    (nò.vi)(tá) 'news'
    (LL)(L)
4. Examining Anchoring

4.1 ANCHOR(σ1), ANCHOR(σstress), SRCS

(27) typology for ANCHOR(σ1), ANCHOR(σstress), SRCS

<table>
<thead>
<tr>
<th>predicted outputs:</th>
<th>1. ANCHOR-σ1, SRCs &gt;&gt; ANCHOR-σstress</th>
</tr>
</thead>
<tbody>
<tr>
<td>unmarked word structure, anchoring to the initial syllable of the base</td>
<td></td>
</tr>
<tr>
<td>ex: Italian André (Andréa), French abrév (abbreviation), Serbo-Croatian</td>
<td></td>
</tr>
<tr>
<td>Svétka (Svetlana)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2. ANCHOR-σstress, SRCs &gt;&gt; ANCHOR-σ1</td>
<td></td>
</tr>
<tr>
<td>predicted outputs:</td>
<td></td>
</tr>
<tr>
<td>unmarked word structure, anchoring to the main-stressed syllable of the base</td>
<td></td>
</tr>
<tr>
<td>ex: French crobe (micróbe), Catalan Fina (Josefina), German Básti (Sebástian)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>3. ANCHOR-σ1, ANCHOR-σstress &gt;&gt; SRCs</td>
<td></td>
</tr>
<tr>
<td>predicted outputs:</td>
<td></td>
</tr>
<tr>
<td>word structure that may be marked, but that preserves both the initial and the</td>
<td></td>
</tr>
<tr>
<td>main-stressed syllable of the base</td>
<td></td>
</tr>
<tr>
<td>ex: cf. below</td>
<td></td>
</tr>
</tbody>
</table>

NB:
- general problem: Anchoring is not systematically investigated in all studies
- to keep matters simple, we assume that the domain for anchoring is always the syllable and not the foot or the segment (cf., e.g., Cabré 1998, Cabré & Kenstowicz 1996 who use the foot as the domain for anchoring in Catalan hypocoristics).
examples: Both anchoring possibilities exist in Italian, French, English

a. Ále Sándra Alessándra
Fránce Céscia Francésca

b. Domí Lodí Dominique Elodie

(c) Pat Trish Patricia Hezekiah
Hez Kye Zabét Elisabet

patterns (26) 1., 2: ... with a twist ...
We predict that if for some reason one anchoring constraint cannot be satisfied, the other will be (if possible)
e.g. French hypocoristics (Nelson 2003: 133ff.; maybe only a tendency)

(29) Domí Dominíque Caró Carolíne
Lodí Elodíe Zabét Elisabet

4.2 Enter C ONTIG: Saving all the prominent material and the template

(32) discontinuous mapping in Spanish hypocoristics
Fénéo Fulgencio
Fíko Federíco
Finda Florínda
Mína Marina

4.3 Enter CONTIG: Saving all the prominent material and the template

(33) ANCHOR(σ₁), ANCHOR(σstress), SRCs >> CONTIGUITY-BT

5. Conclusion/Outlook
We looked at:
• truncations: templates and anchoring
• theoretical machinery: GTT, anchoring constraints
• database: so far, 76 patterns, 22 languages
the main findings:
• GTT predicts more than has hitherto been in the focus of pertinent studies.
  o different foot templates (maximal or wellformed?)
  o different numbers of feet (one or many?)
  o if interacting with syllable wellformedness: subminimal foot templates
  => We find these patterns in our sample!
• Anchoring constraints predict more than has hitherto been in the focus of most pertinent studies.
  o the existence of ‘anchoring alternatives’
  o anchoring as a force generating atemplatic truncation
  o discontinuous mapping
  => We find most of these patterns in our sample!
• In combination, GTT and anchoring constraints make yet more interesting predictions
  o ‘well-behaved’ truncation (good template, good anchoring): is only one of several possibilities
  o ‘not-so-well-behaved truncation’: dominance of anchoring constraints over SRCs may lead to
    ▪ variable word size
    ▪ marked word structure (iambic truncation in trochaic languages)
  => We find these patterns in our sample!

outlook: lots of things for us to do
This paper was only the first step. In the future, we plan to
• collect more patterns
  in order to be able to
  • develop a more detailed typology of attested ranking permutations among SRCs
  • get a clearer picture of anchoring generalisations (e.g.: What about base-final anchoring? Are there languages in which prominence anchoring does not matter?)

References


Truncation patterns quoted in the handout and their sources:
Dutch: van de Vijver (1997)
English: Lappe (2003, 2005)
German: Wiese (2001), Alber (in press)
Italian: Alber (in press)
Zuñi: Weeda (1992)