

150 Years after Dillmann's Lexicon:
Perspectives and Challenges of Gəʿəz Lexicography

Supplement to *Aethiopica*.
International Journal of Ethiopian
and Eritrean Studies

5

Edited in the Asien-Afrika-Institut
Abteilung für Afrikanistik und Äthiopistik
Hiob Ludolf Zentrum für Äthiopistik
der Universität Hamburg

Series Editor: Alessandro Bausi
in cooperation with Bairu Tafla, Ulrich Braukämper,
Ludwig Gerhardt, Hilke Meyer-Bahlburg

2016

Harrassowitz Verlag · Wiesbaden

150 Years after Dillmann's Lexicon:
Perspectives and Challenges
of Gə'əz Studies

Edited by
Alessandro Bausi
with assistance from
Eugenia Sokolinski

2016

Harrassowitz Verlag · Wiesbaden

The publication of this volume was supported by the European Union Seventh Framework Programme IDEAS (FP7/2007-2013) ERC grant agreement 338756 (TraCES).

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the internet at <http://dnb.dnb.de>.

For further information about our publishing program consult our website <http://www.harrassowitz-verlag.de>

© Otto Harrassowitz GmbH & Co. KG, Wiesbaden 2016

This work, including all of its parts, is protected by copyright. Any use beyond the limits of copyright law without the permission of the publisher is forbidden and subject to penalty. This applies particularly to reproductions, translations, microfilms and storage and processing in electronic systems.

Printed on permanent/durable paper.

Typesetting, copy editing, index: Eugenia Sokolinski (Hamburg).

Printing and binding: Memminger MedienCentrum AG

Printed in Germany

ISSN 2196-7180

ISBN 978-3-447-10783-9

Table of Contents

Preface	vii
Notes to the reader	xi
Introduction. <i>150 Years After Dillmann's Lexicon</i> (A. BAUSI)	3
Chapter 1. Research in Gəʿəz linguistics	11
<i>The TraCES project and Gəʿəz studies</i> (E. SOKOLINSKI)	13
<i>A part of speech tag set for Ancient Ethiopic</i> (S. HUMMEL, W. DICKHUT)	17
<i>Bringing Gəʿəz into the digital era: computational tools for</i> <i>processing Classical Ethiopic</i> (C. VERTAN)	31
<i>On editing and normalizing Ethiopic texts</i> (A. BAUSI)	43
<i>Some problems of transcribing Geez</i> (M. BULAKH)	103
Chapter 2. Language contact	139
<i>Sabaic loanwords in Gəʿəz and borrowings from Gəʿəz into</i> <i>Middle Sabaic</i> (S. FRANTSOZOFF)	141
<i>Nasal infix as index of Semitic loanwords borrowed through</i> <i>the Greek</i> (A. SOLDATI)	149
<i>New Gəʿəz word forms from Arabic-Ethiopic translation literature.</i> <i>Suggestions for lexical entries and their meanings, as demonstrated</i> <i>from Secundus the Silent Philosopher</i> (M. HEIDE)	173
Chapter 3. Gəʿəz lexicography in comparison	183
<i>Beyond Dillmann's Lexicon – Towards digital lexicography:</i> <i>Lessons from Syriac</i> (A. ELLWARDT)	185
<i>Sergew Hable Selassies Fragment eines Gəʿəz-Belegstellenlexikons und</i> <i>Abraham Johannes Drewes' Glossare zum Recueil des inscriptions de</i> <i>l'Éthiopie. Zwei unveröffentlichte Beiträge zur äthiopischen Lexiko-</i> <i>graphie und deren Bewertung und Lehren für die heutige informations-</i> <i>technisch aufgerüstete Äthiopistik</i> (M. KROPP)	201
<i>The use of Arabic in Gəʿəz lexicography: from Dillmann to Leslau</i> <i>and beyond</i> (S. WENINGER)	219
Index	233

A part of speech tag set for Ancient Ethiopic*

SUSANNE HUMMEL and WOLFGANG DICKHUT, Universität Hamburg

In this article, we would like to share the results obtained to date in the course of establishing a comprehensive part of speech (PoS) tag set to be used for the linguistic annotation of Gəʿəz texts. Questions regarding the editions used, including necessary adjustments or corrections, will not be considered here.¹

General presentation of the tag set

The annotation used here presupposes a tokenization of complex graphic units which cannot be done on the Ethiopic script (*fidal*). By graphic unit we mean the sequence of characters separated by a word divider (⌘) or a punctuation mark (⌘). Due to the syllabic script, the process of tokenization can only be done on the transliteration.² The GeTa tool provides an automatically generated transcription, in which the elimination of the shwa (ə) and the gemination of consonants have to be corrected manually, e.g.:

Mt 1:18	ተረክበት ፡ ዕንስተ ፡ እሙንፈስ ፡ ቅዱስ ።
automatic transcription:	<i>tarakəbat dənəsətə ʿəmanəfas qədus.</i>
corrected transliteration:	<i>tarakbat dənəstə ʿəmmənfəs qəddus.</i>
tokenized:	<i>tarakbat dənəstə ʿəmm-manfas qəddus.</i>

‘she was found with child of the Holy Ghost’³

We annotate the transliterated and tokenized texts in two ways: firstly, by way of a basic annotation (e.g. marking named entities)—which is not the topic of our paper; secondly, by using a fine-grained tag set, resulting in a detailed linguistic annotation.

The criteria for annotating are, of course, mainly, but not exclusively, morphological. For example, the morphologically identical forms of the ac-

* The research leading to these results has received funding from the European Research Council under the EU Seventh Framework Programme, grant agreement no. 322849.

1 Variant readings, misprints, etc. are accounted for in the process annotation by using the commenting function of the GeTa tool. On the GeTa annotation tool, see Cristina Vertan’s essay in this volume.

2 The tokenization cannot be applied on the Ethiopic script. For example, the graphic unit ልደቱ ፡ *lədatu* is composed of three letters corresponding to the three syllables ል *lə*, ደ *da* and ቱ *tu*, but it has to be tokenized as *lədat-u* ‘his nativity’.

3 Note that when the proclitically used preposition ʿəm is followed by a word beginning with *m*, like *manfas*, we will often find only a single ‘ሙ’ (*m[a]*) written, which is then geminated (*mm[a]*). The examples are drawn from the so-called A-text of the Gospel of Matthew, as edited in Zuurmond 2001.

Table 1. Parts of speech and relevant features.

Class	Subclass		Tag	Description	Features
Nominals	Nouns <u>N</u>	1	NCom	Common Noun	Gender, Number, Case, State
		2	NProp	Proper Name	Gender, Number, Case
	Pronouns <u>P</u>	3	PPers	Independent Personal Pronoun	Person, Gender, Number, Case አኅ ፣ አንተ ፣ አንቲ ፣ ውሉቱ ፣ ውሉት ፣ ይእቲ ፣ / ይእት ፣ ...
		4	PSuff	Pronominal Suffix	Person, Gender, Number +የ/ +ኒ, +ከ, +ከ, +ሀ, +ሃ ...
		5	PSub	Subject Pronoun Base	no features ለለ, +PSuff (but 1c.sg.: ለለ,የ ፣ or ለለ,የ ፣ or ለለ,የ ፣)
		6	PObj	Object Pronoun Base	no features ከ,የ+PSuff
		7	PPoss	Possessive Pronoun Base	Gender, Number m.sg. ከ,አ+PSuff , f.sg. እንቲአ+PSuff , c.pl. እለ,አ+PSuff
		8	PDem	Demonstrative Pronoun	Gender, Number, Case • near: m.sg. ዝ+/ዘ+ or ዝንቲ ፣/ዘንቲ ፣ , f.sg. ዛ+ or ዛቲ ፣/ዛት ፣ m.pl: እሉ ፣ or እሉንቲ ፣/እሉንት ፣ , f.pl. እላ ፣ (እሉን ፣) or እላንቲ ፣ (እላንቲ ፣) / እላንት ፣ • far: m.sg. ዝኩ ፣/ዝኩ ፣ or ዝኩቲ ፣/ዝኩት ፣ , f.sg. እንትኩ ፣/እንትኩ ፣ or እንታኩቲ ፣/እንታኩት ፣ , c.pl. እልኩ ፣ or እልኩቲ ፣ (እልኩቲ ፣) / እልኩት ፣ (እልኩት ፣) or እላኩ ፣/እላኩት ፣
		9	PRel	Relative Pronoun	Gender, Number m.sg. ዘ+ , f.sg. እንት ፣ , c.pl. እለ ፣
		10	PInt	Interrogative Pronoun	Number, Case መኑ ፣/መኑ ፣ <i>who?</i> , ምንት ፣/ምንት ፣ <i>what?</i> , sg. አይ ፣/አየ ፣ <i>which?</i> , pl. አ,የት ፣/አ,የት ፣
		11	PTot	Pronoun of Totality Base	Case ኩሉ+PSuff <i>all, everything</i> ኩሉንታ+PSuff <i>whole, in the whole being</i>
		12	PSol	Pronoun of Solitude Base	Case ባሕቲ፣ት+PSuff <i>alone</i>
	Numerals <u>Num</u>	13	NumCa	Cardinal Numeral	Gender, Number, Case, State, Logogram
		14	NumOr	Ordinal Numeral	Gender, Number, Case, State, Logogram
	Verbs <u>V</u>	15	V	Verb	Person, Gender, Number, Tense/Aspect/Mood For Infinitive: Case, State
Existentials <u>Ex</u>	Existentials	16	ExAff	Existentials Affirmative Base	no features • በ+PSuff <i>there is</i> (በየ ፣ በከ ፣ በከ, ፣ በ ፣ or ቦ፣ ፣ ቦ ፣ or ቦ፣ ፣ ...)
		17	ExNeg	Existentials Negative Base	no features • አልበ+PSuff : (አልብየ ፣ አልብከ ፣ አልብከ, ፣ አልቦ ፣ or አልቦ፣ ፣ አልባ ፣ or አልባ፣ ፣ ...)

Class	Subclass		Tag	Description	Features
Particles	Adverbs △	18	AInt	Interrogative Adverb	no features አይቲ፣ ፣ where? ማእዜ ፣ when? እር፣ ፣ how? ሚመጠን ፣ how many?
		19	A	Other Adverb	no features • location: ህየ ፣ there, አፍኦ ፣ outside, ዝየ ፣ here, ፀጋም ፣ left • time: ቀዲሙ ፣ before, ካፅበ ፣ again, ዮም ፣ today, ይእዜ ፣ now, ኒሰም ፣ tomorrow • others: መጠን ፣ as large as, ቀዳሚ ፣ at first, ባሕቲ only, አማን ፣ truly, ግሙራ ፣ entirely
	Preposition Prep	20	Prep	Preposition	State (nominal and pronominal state) • enclitic: ለ+ for, በ+ with, እም+ from • independent: ላዕለ ፣ above, ምስለ ፣ with, ማእከለ ፣ between, ቅድመ ፣ in front, ንበ ፣ to • compound: በኅበ ፣ at, ዘእንበለ ፣ without, በእንተ ፣ concerning, እንበይነ ፣ for the sake of • comparative: እም ፣ than, ከመ ፣ like
	Conjunction Conj	21	Conj	Conjunction	no features • ባሕቲ ፣ yet, ሶበ ፣ when, ንበ ፣ where, በኅበ ፣ (there,) where, አላ ፣ but rather, እም ፣ since, አመ ፣ when, እመ ፣ if, አሰመ ፣ because, እስከ ፣ until, እንዘ ፣ while, አው ፣ or, ከመ ፣ (so) that, ወ+ and, ዘ ፣ that, ዳእሙ ፣ but • ለ+subjunctive so that
	Interjection Int	22	Int	Interjection	State (nominal and pronominal state) • to express woe!: አሌ ፣ (+ለ+PSuff: አሌልየ ፣ or አሌ ፣ ለተ ፣), ወይ ፣ • ለ+ ob!, አህ ፣ ab!, እንቋፅ ፣ ba!, ዩ ፣ or ዮ ፣ alas!, ጸጥ ፣ quiet! • ባሕ ፣ bail! (also +PSuff 2m/f: ባሕከ ፣ ባሕኪ ፣)
	Further Particles	23	PartAcc	Accusative Particle	no features +ሃ (suffixed only to proper names)
		24	PartAff	Affirmative Particle	no features እው ፣ , +ን yes indeed!, certainly!
		25	PartDei	Deictic Imperative Particle	Person, Gender, Number m.sg. ነጻ ፣ (ንጻ ፣), f.sg. ንጻ ፣ m.pl. ንዑ ፣, f.pl. ነጻ ፣
		26	PartInt	Interrogative Particle	no features +ሀ, +ኑ
		27	PartNeg	Negative Particle (Base)	no features አ+, አልበ ፣ (= PartNeg+PSuff:3m.sg), አንበ, +PSuff: እንብየ ፣, አኮ ፣
		28	PartPres	Presentational Particle Base	no features Behold! ኅ+PSuff:3m.sg (ኅሁ ፣), ኅው+PSuff:3f.sg (ኅዋ ፣), ነይ +PSuff (not second person), እን+PSuff
		29	PartQuot	Quotative Particle	no features +አ
		30	PartVoc	Vocative Particle	no features አ+
31		PartOth	Other Particles	no features +ሂ and also, +ሰ even, +ሶ doch!, +ኔ even, +ኑ to, እስከ ፣ O doch!, እንጋ ፣ so then!, +ኪ now, ክመ ፣ just	
Foreign Material	32	ForMat		For example ሰገድ sagad (element in proper names such as ሳናጎግ Sagad, regnal name of King Galawdewos, r. 1540–1559)	
Punctuation	33	Punct			

cusative and the construct state have to be separated into (a) accusative, (b) construct state, (c) accusative and construct state, e.g.:

Mt 2:11	ወበዊአሙ ፡ ቤተ ፡	<i>wa-bawī'omu beta</i>
	‘and when they came into the house’	acc + absSt
Mt 4:5	ዲበ ፡ ማእከት ፡ ቤተ ፡ መቅደስ ፡	<i>diba māzənta beta maqdas</i>
	‘on a pinnacle of the temple’	nom + conSt
Mt 8:14	ወዐአ ፡ ኢየሱስ ፡ ቤተ ፡ ጴጥሮስ ፡	<i>wa-bo'a 'iyasus beta petros</i>
	‘and Jesus came into Peter’s house’	acc + conSt

In the same way, we mark whether and in which way the gender of a noun is determined by the context (i.e. by syntax, not morphologically), for example by the corresponding verb (see below).

Apart from the present authors’ own experience of reading Ethiopic texts, the proximate sources of the tag set were (1) other tag sets (for general considerations)⁴ and (2) the standard reference grammars, namely those of Dillmann and Tropper.⁵ On that basis we established a preliminary tag set which we then improved and refined during the first stages of the annotation process.

To disambiguate morphologically identical forms which have different meanings and are thus different PoS we take syntax and semantics into account. Disambiguation is necessary already at the level of tokenization, as in **ሞቲ ፡** *motu*, ‘they died’ or *mot-u* ‘his death’, depending on context. The same is true of ambiguous forms like **ገብረ ፡** *gabra* with the function of a verb ‘he worked’ or of a common noun in the accusative or the construct state ‘slave, servant’; the graphic unit **አም ፡** can be read as *ʾamm* ‘mother’ or *ʾam* ‘from’. Other examples of similar phenomena:

- አዳም ፡** *ʾaddām*, proper name (‘Adam’) or common noun (‘(a) pleasant one’);
- እንተ ፡** *ʾənta*, relative pronoun fem. sing. or preposition (‘through’, ‘towards’);
- ኻበ ፡** *ḥaba*, conjunction (‘where’) or preposition (‘towards’);
- ከመ ፡** *kama*, conjunction (‘as’) or preposition (‘as’, ‘like’).

The same is true of many composed expressions, for example:

	በከመ ፡ <i>ba-kama</i> , Prep-Conj or Prep-Prep, as in
Mt 1:24	ገብረ ፡ በከመ ፡ አዘዞ ፡ መልአክ ፡ እግዚአ ፡ ብሔር ፡
	<i>gabra ba-kama</i> (Prep-Conj) <i>ʾazzaz-o mal'aka ʾəgzī'a bəḥer</i>
	‘[Joseph] did as the angel of the Lord had bidden him’

4 For example, earlier versions of Zeldes and Schroeder 2016.

5 Dillmann 1899, 1907; Tropper 2002.

Mt 16:27 **ይፈጅዮ ፡ ለከሉ ፡ በከመ ፡ ምግባሩ ።**

yafadday-o la-k wall-u ba-kama (Prep-Prep) *məḡbār-u*

‘he shall reward every man according to his works’

Note also that adverbial expressions which are the accusative forms of a common noun are tagged accordingly, for example:

Mt 1:19 **አብደረ ፡ ይኅድጋ ፡ ጽምግታ ፡**

ʾabdara yəḥdāg-ā šammīta (NCom + acc)

‘[Joseph] was minded to put her away privily’

As a general rule, all the information contained in the annotation should be based on internal evidence, without enforcing predetermined rules. In cases where combinations of forms have received a relatively stable meaning of their own, we have accepted Dillmann’s *Lexicon* as a guide in deciding whether the components should be annotated separately or an entire compound should be treated as one unit. The same applies to infinitives used (and categorized by Dillmann) as common nouns, e.g. **አእምሮ ፡** *ʾaʾmāro* ‘knowledge’.⁶ It goes without saying that the authority given to Dillmann in this way is only relative, and his decisions can be modified if appropriate.

Another case in which a reasonable predetermined rule can be applied is the distinction between the preposition *ba-* + suffixes, and the same combination expressing the ‘affirmative existential’ (‘there is’), which very frequently gives a possessive meaning (‘he has’), for example:

Mt 27:42 **ወንእመን ፡ ቦቲ ፡**

‘and we shall believe in him’

wa-naʾman b-ottu

Prep-PSuff.3m.sg

Mt 12:11 **ዘቦቲ ፡ አሐተ ፡ በግዕተ ፡**

‘who shall have one sheep’

za-b-ottu ʾahata baggāʿta

ExAff-PSuff.3m.sg

The complete set of PoS with their correspondent features is represented in Table 1. In the following we would like to focus on the more complex annotation of nouns.⁷

Annotation of nouns

We divide nouns into common nouns and proper names. The tag set does not distinguish between nouns and adjectives since they have the same mor-

6 Other instances include, for example, **ተነብዮ ፡** (vel **ተነብዮት ፡**) ‘charisma vel munus prophetiae, prophetia, vaticinium’; **ሐቂይ ፡** (vel **ሐቅይ ፡**) ‘stridor, frendor [dentium]’; **አምልኮ ፡** vel **አምልኮት ፡** ‘εὐσέβεια religio, pietas [erga Deum vel numen divinum]’; **ተሰእሎት ፡** ‘quaestio, interrogatio’; **ተጋሥሥ ፡** ‘ζητήσεις quaestio, controversia’; **ተወቆ ፡** ‘αἴσθησις cognitio, scientia, perspicientia’; **አጣዕዎ ፡** ‘idololatria, magia’.

7 The examples given are drawn from the *Kəbra nagašt*, as edited in Bezold 1905.

phological properties. For example, the forms **አያው** : *ḥayāw* ‘living’ or **ምድራዊ** : *madrāwi* ‘earthly’ can have both functions, namely of an adjective and of a noun. Nevertheless, since nouns in the function of adjectives play a role in the annotation process (whenever syntax has to be considered), the term adjective is also used here to describe the annotation scheme.

While, as a rule, common nouns are morphologically built up according to the Semitic root-pattern system, many proper names are of foreign origin and thus lie outside this system. Besides, Ethiopian personal names generally consist of several, two or three, common nouns, often in a genitive construction, like **ተክለ ገይማኖት** : *Takla Hāymānot* or **ገብረ መንፈስ ቅዱስ** : *Gabra Manfas Qəddus*. Each element of an Ethiopian personal name is annotated as an (inflected) common noun on the morphological annotation level (with each element tagged as ‘part of a proper name’) and linked to the respective lemma in the dictionary, while the entire name is annotated as one named entity at the basic annotation level and linked to a prosopographic database.

We assign the following features to nouns:

Gender	Masculine <u>m</u>	Feminine <u>f</u>	Unmarked <u>unm</u>
	Nature or Pattern and/or Syntax		
Number	Singular <u>sg</u>	Plural <u>pIn, pEx, pp</u>	Unmarked <u>unm</u>
	Pattern: sg pIn pEx pp unm	Syntax: sg pl unm	
Case	Nominative <u>nom</u>	Accusative <u>acc</u>	Vocative <u>voc</u>
State	Absolute State <u>absSt</u>	Construct State <u>conSt</u>	Pronominal State <u>pronSt</u>

The feature of ‘state’ does not apply to proper names. In all examples below, abbreviated forms of features are given in the following sequence: Gender-Number-Case-State.

Gender

The nominal feature of ‘gender’, with its three values ‘masculine’, ‘feminine’ and ‘unmarked’, is one of the most complex within the whole tag set. The values ‘masculine’ and ‘feminine’ are specified according to different criteria by which a noun’s gender is marked. We annotate whether the gender is marked by morphology (i.e. noun formation through a certain ‘pattern’) or

by semantics ('nature'), and/or whether it is identifiable within a clause or sentence by syntactic agreement ('syntax'). The specified value 'syntax' may be applied alone or together with either 'pattern' or 'nature'. We formulated several rules (i–ix) to consistently apply these further specifications.

Pattern or Nature

i) The gender of a noun is determined by pattern or nature if there exists a binary opposition between a masculine and a feminine nominal form, either on morphological or on lexical grounds. Thus, we distinguish binary pairs having the same root (determined by pattern) and pairs existing only on the lexical level, originating from different roots (determined by nature). Typical examples for binary pairs by pattern are ብእሲ : *bə'si* 'man'—ብእሲት : *bə'sit* 'woman', ቅዱስ : *qəddus* 'saint(ly)' masculine—ቅድስት : *qəddəst* 'saint(ly)' feminine or ጸሊም : *šallim* 'black' masculine—ጸላም : *šallām* 'black' feminine; patterns with morpho-phonological alternations like ወልድ : *wald* 'son'—ወለት : *walatt* 'daughter' are also distinguished. If the gender of a noun is determined by pattern, we ignore a possible natural gender in order to distinguish strictly between the two specifications.

Consequently, for example when an apparently feminine form (from the point of view of Semitic noun formation) is employed for abstract nouns (e.g. ትውልድ : *təwələdd* 'generation'), we do not treat it as feminine by pattern, since there is no opposition between a masculine and a feminine noun form. This is in line with the scope of the linguistic annotation, i.e. drawing data from the texts and not limiting the annotation to general principles of noun formation.

ii) Some nouns have only one plural form for both masculine and feminine. In such cases, the gender is not marked in the plural form (i.e. 'unm. '), though it is marked by 'pattern' (i.e. ^P) in the singular form; for example አድግ : *ʾadg* (m^P.sg^P), 'donkey' masculine, አድግት : *ʾadgət* (f^P.sg^P) 'donkey' feminine and አእዱግ : *ʾaʾdug* (unm.pln) 'donkeys'.

iii) Classic examples of binary pairs defined by nature and not by pattern are አብ : *ʾab* 'father'—እም : *ʾamm* 'mother' or ገብር : *gabr* 'servant'—አመት : *ʾamat* 'maid', whereas nouns like መልአክ : *malʾak* 'messenger', 'angel', are considered as unmarked for gender by pattern (but may be determined by syntactic agreement).

iv) The gender of most personal names is determined by nature as is obvious in the examples, ማርያም : *māryām* 'Mary' and ኢየሱስ : *ʾiyasus* 'Jesus'.

Syntactic agreement

While the specifications of pattern and nature are mutually exclusive (if pattern applies, nature is ignored), the specific value 'syntax' is available for

both. We apply this specific value if the gender can be deduced from agreement with verbs, pronouns or adjectives within the same clause or sentence as the following examples demonstrate:

1) **ወኩሉ ፡ ግብሩ ፡**

‘and all his doing’

wa-k wəll-u

Conj-PTot:nom-PSuff:3m.sg

gəbr-u

NCom:m^s.sg^{PS}.nom.pronSt-PSuff:3m.sg

2) **ትንግሮሙ ፡ ለደቂቅክ ፡**

‘so that you shall tell your children’

təngərr-omu

V:subj.2m.sg-PSuff:3m.pl

la-daqqiqa-ka

Prep:nomSt-NCom:m^{PS}.sg^P.pl^S.nom.pronSt-PSuff:2m.sg

3) **ኅለይኩ ፡ በእንተዝ ፡ ነገር ፡**

‘I thought about this matter’

hallayku

V:perf.1c.sg

ba’ənta-zə

Prep:nomSt-PDem:m.sg

nagar

NCom:m^s.sg^{PS}.nom.absSt

Taking syntactic agreement into account significantly adds to the complexity of the whole annotation process. We therefore established special rules to facilitate a homogenous annotation.

v) We do not consider the relative pronoun **ዘ ፡ za** (masculine, singular) or the independent personal pronoun **ውእቱ ፡ wə’ətu** (third person, masculine, singular) in determining syntactic agreement since both are commonly used for all genders and numbers.

vi) A pronominal suffix attached to the existential base refers to the possessor and thus marks its gender when the possessor is mentioned, as in the sentence:

4) **ንሕነ ፡ ሕዝብ ፡ ዘአልቦ ፡ ሕገ ፡**

‘we are people without law’

nəhna

PPers:1c.pl

həzb

NCom:m^s.sg^{PS}.nom.absSt

za-’alb-o

PRel:m.sg-ExNeg-PSuff:3m.sg

həgga

NCom:unm^(gender).sg^P.acc.absSt

In existential clauses, a pronominal suffix attached to the existential base does not indicate syntactic agreement:

5) **ወአልቦ ፡ ሐሰት ፡ ውስተ ፡ ቃልክ ፡**

‘there is no lie in your word’

wa-’alb-o

Conj-ExNeg-PSuff:3m.sg

hasat

NCom:unm^(gender).unm^(number).nom.absSt

wəsta

Prep:nomSt

qālə-ka

NCom:m^s.sg^P.nom.pronSt-PSuff:2m.sg

vii) Only the immediate context is significant when identifying syntactic agreement. If a series of elements stands in agreement with one (other) single element, we assume that the determiner determines only that element which is closest to the subject:

6) ወእዘዎ፡ እግዚአብሔር ፡ ለሙሴ ፡ ወለአሮን ፡

‘And God gave order to Moses and Aaron’

wa-²azzaz-omu ²*əgzī* *ābəḥer* *la-muse*

Conj-V:perf.3m.sg-PSuff:3m.pl NProp:m^{PS}.sg^{PS}.nom Prep:nomSt-NProp:m^{NS}.unm^(number).nom

wa-la-²aron

Conj-Prep:nomSt-NProp:m^N.unm^(number).nom

In example (6), the gender of the proper name *muse* (masculine by nature) is treated as syntactically marked, while the gender of the name ²*aron* is considered as determined only by nature.

7) ወኅልቀ ፡ ምስሌሆሙ ፡ ከኑሉ ፡ እንሰሳ ፡ ወአራዊት ፡

‘And all cattle and animals perished with them’

wa-ḥalqa *məsle-homu* *kəwəll-u*

Conj-V:perf.3m.sg Prep:pronSt-PSuff:3m.pl PTot:nom.-PSuff:3m.sg

²*ənsəsā* *wa-²arāwīt*

NCom:m^S.sg^{PS}.nom.absSt Conj-NCom:unm^(gender).pIn^P.nom.absSt

In example (7), only the first noun, ²*ənsəsā*, is treated as syntactically marked.

In the synthetic genitive construction, only the head in the construct state is marked for gender:

8) በደብረ ፡ ጽዮን ፡ ዘኅደርክ ፡ ውስቴታ ፡

‘on Mount Zion, where you have dwelt’

ba-dabra *ṣəyon*

Prep:nomSt-NCom:f^S.sg^{PS}.nom.conSt NProp:unm^(gender).unm^(number).nom

za-ḥadarka *wəstet-ā*

PRel:m.sg-V:perf.2m.sg Prep:pronSt-PSuff:3f.sg

viii) Apposition does not pre-determine the gender and is not used to determine and indicate gender by syntax. In the phrase ወክብራ ፡ ለጽዮን ፡ ታቦተ ፡ ሕጉ ፡ ለእግዚአብሔር ፡ *wa-kəbr-ā la-ṣəyon tābota ḥəgg-u la-²əgzī ābəḥer* ‘and the glory of Zion, ark of the covenant of the law of God’, we cannot draw any conclusions as to the gender of the noun *tābot*, whereas the proper name *ṣəyon* (with no natural gender) is referred to with a feminine suffixed pronoun attached to the common noun *kəbr*.

Cases like the following one are no exception. In ወበወልዱ ፡ ኢየሱስ ፡ *wa-ba-wald-u ²iyasus* ‘and with his son Jesus’, we do not mark the syntactic gen-

der for ‘Jesus’ (masculine by nature) even if in this precise case *ʿiyasus* is in apposition to *wald* ‘son’, which in turn is masculine by pattern.

ix) The rule for apposition also applies to a nominal predicate because a masculine subject can have a female predicate (and vice versa). The nominal predicate is not marked syntactically; in the phrase **ወኮነ ስቅለ** : *wa-kona baqla* ‘and the mule came into being’, the gender is not marked. However, the predicate adjective is a different case since it modifies the subject; in the clause **ወደቂቀ ስናኣንሰ ጂ ጽኑዓን** : *wa-daqiqa kanāʿānə-ssa 7 šənuʿān* ‘and the sons of Canaan were seven strong (men)’, the collective noun *daqiq* is not only marked by pattern (masculine) but also through syntactic agreement with the noun *šənuʿān* (masculine by pattern).

Number

Gəʿəz has both singular and plural. The plural is formed either by the attachment of the suffixes *-ān* (masculine) or *-āt* (feminine) to the singular and is referred to as ‘external plural’, or by a specific pattern change which is referred to as ‘internal plural’. The suffix *-āt* can also be attached to an already formed plural, i.e. ‘plural of plural’; for example **ሊቃ ሊቃን** : *liq* (sg.)—**ሊቃን** : *liqān* (pEx)—**ሊቃናት** : *liqānāt* (pp) or **ለብስ** : *ləbs* (sg.)—**ለልባስ** : *ʿalbās* (pIn)—**ለልባሳት** : *ʿalbāsāt* (pp).

We further specify whether number is marked morphologically (‘pattern’) and/or syntactically (‘syntax’). Nouns are treated as marked by ‘pattern’ either when a singular or a plural form is attested or when the pattern is able to produce a plural form. Otherwise number is unmarked by pattern. Independently of the morphological form we assign number according to syntactic agreement. The following examples—concerning nouns that can behave as collective nouns—demonstrate our approach:

9) **ወሰበ ርእዩ ሰቂቀ ገዳዳ ለስራኤል** :

‘and when the sons of Israel’s dignitaries saw’

<i>wa-soba</i>	<i>rəʿyu</i>	<i>daqiqqa</i>
Conj-Conj	V:perf.3m.pl	NCom:m ^{PS} .sg ^P .pl ^S .nom.conSt
<i>ḥayyālāna</i>	<i>ʿIsrāʿel</i>	
NCom:m ^P .pEx ^P .nom.conSt	NProp:unm ^(gender) .unm ^(number) .nom	

10) **ወበዝነ ለክዮሙ ለደቂቀ ሰዓን** :

‘and the wickedness of the children of Cain increased’

<i>wa-bazḥa</i>	<i>ʿakay-omu</i>	<i>la-daqiqqa</i>
Conj-V:perf.3m.sg	NCom:m ^S .sg ^P -PSuff:3m.pl	Prep:nomSt-NCom:m ^{SP} .sg ^P .pl ^F .nom.conSt
<i>qāyan</i>		
NProp:m ^N .unm ^(number) .nom		

11) አመ ፣ ቀተለ ፣ እግዚአብሔር ፣ በኮሮሙ ፣ ለግብጽ ፣

‘when God killed the firstborn of Egypt’

^o ama	qatala	^o agzi ^o ābāber
Conj	V:perf.3m.sg	NProp:m ^{PS} .sg ^{PS} .nom
bak ምጥ-omu		la-gəbṣ
NCom:unm ^(gender) .sg ^P .acc.pronSt-PSuff:3m.pl		Prep:nomSt-NProp:m ^S .pl ^S .nom

In many cases there is no syntactic reference to number in a clause or sentence, as in the following examples:

12) ላዕለ ፣ ደቂቀ ፣ አዳም ።

‘against the sons of Adam’

lā ^o la	daqiqā	^o addām
Prep:nomSt	NCom:m ^P .sg ^P .nom.conSt	NProp:m ^N .unm ^(number) .nom

13) ወቦሙ ፣ ይመውእ ፣ ፀሮ ።

‘and with them he defeated his enemy’

wa-b-omu	yəməwəwə ^o	ḍar-o
Conj-Prep:pronSt-PSuff:3m.pl	V:impf.3m.sg	NCom:unm ^(gender) .sg ^P .acc.pronSt-PSuff:3m.sg

Nouns that have no morphological plural (like ሰብእ ፣ *sab^o* ‘people’) are annotated as unmarked by pattern, and their number can be deduced from syntactic agreement.

14) እከውን ፣ ሰብእ ፣

‘I will become man’

^o akawwən	sab ^o a
V:impf.1c.sg	NCom:unm ^(gender) .unm ^(number) .acc.absSt

15) ወይስድድዎ ፣ ለሰብእ ፣ ብሔረ ፣ ኢትዮጵያ ፣

‘so that they pursue the people of Ethiopia’

wa-yəsdədəw-o	la-sab ^o a
Conj-V:subj.3m.pl-PSuff:3m.sg	Prep:nomSt-NCom:m ^S .sg ^S .nom.conSt
bəḥera	^o ityopyā
NCom:unm ^(gender) .sg ^P .nom.conSt	NProp:unm ^(gender) .unm ^(number) .nom

16) በእንተ ፣ ዙተፈሥሐ ፣ ሰብእ ፣ ኢትዮጵያ ።

‘how the people of Ethiopia rejoiced’

bā ^o anta	za-tafāsśəḥu
Prep:nomSt	PRel:m.sg-V:perf.3m.pl
sab ^o a	^o ityopyā
NCom:m ^S .pl ^S .nom.conSt	NProp:unm ^(gender) .unm ^(number) .nom

State and Case

Gə^cəz has three cases (nominative, accusative and vocative) and three states (absolute state, construct state and pronominal state) for which nouns may

or may not undergo vowel changes. The suffix *-o* marks the vocative. The suffix *-a* marks the accusative as well as the construct state and appears in both singular and plural. Nouns ending with *-i* turn into the ending *-e* in both the accusative and the construct state, as in ብእሲ ፡ *bəʾsi* ‘man’ – ብእሴ ፡ *bəʾse*. Nouns ending with *-e*, *-o*, *-ā* have identical forms in the nominative/absolute state and the accusative/construct state, as ጸጌ ፡ *ṣəge* ‘flower’, እእምሮ ፡ *ʾaʾmāro* ‘knowledge’, ሐራ ፡ *ḥarā* ‘army’. For such nouns case and state opposition does not surface. Therefore, many forms are morphologically ambiguous and can only be analysed by taking syntax and semantics into consideration.

Consequently, we assign the features ‘case’ and ‘state’ when a noun is not overtly marked for accusative and/or construct state; we do so on the basis of syntax and semantics:

17) እስመ ፡ አረሩ ፡ ፍሬ ፡ መቅዋፍቶሙ ።

‘for they reaped the fruit of their punishment’

ʾasma ʾararu fəre maqṣaft-omu

Conj V:perf.3m.pl NCom:unm^(gender).sg^P.acc.conSt NCom:unm^(gender).sg^P.nom.pronSt-PSuff:3m.pl

18) ከመ ፡ አይትወደዉ ፡ ደወለ ፡ ቢጸሙ ፡

‘that they shall not transgress the boundaries of their neighbours’

kama ʾi-γətʿ adawu

Conj PartNeg-V:subj.3m.pl

biṣ-omu

dawala

NCom:unm^(gender).sg^P.acc.conSt

NCom:unm^(gender).sg^P.nom.pronSt-PSuff:3m.pl

19) ከመ ፡ ትንበሩ ፡ አንተ ፡ ወቤተ ፡ አቡክ ፡

‘so that you shall remain, you and your father’s house’

kama tənbaru ʾanta

Conj V:subj.2m.pl PPer:2m.sg

wa-beta

ʾabu-ka

Conj-NCom:unm^(gender).sg^P.nom.conSt NCom:m^N.sg^P.nom.pronSt-PSuff:2m.sg

20) ፈቀደ ፡ ከመ ፡ ይሕንጽ ፡ ቤተ ፡ እግዚአብሔር ፡

‘he wished to build the House of God’

faqada kama γəḥnəṣ beta ʾəgziʾābəher

V:perf.3m.sg Conj V:subj.3m.sg NCom:unm^(gender).sg^P.acc.conSt NProp:m^P.sg^P.nom

The same rule applies with the pronominal state. A noun stands in the pronominal state when it attaches a pronominal suffix (see also examples above). If a noun takes a pronominal suffix, its case can only be deduced from second person suffixes or from the suffix of the third person masculine, as in ‘his house’ ቤቱ ፡ *bet-u* (nominative) or ቤቱ ፡ *bet-o* (accusative). Nouns with all other suffixes (first person and all the other third person suffixes) are not overtly marked.

Proper names, which cannot stand in the construct state, either have the same form in the nominative and the accusative (here in accusative: **ወበጽሐት ፡ ኢየሩሳሌም ፡** *wa-baṣḥat ʾiyarusālem* ‘and she reached Jerusalem’), or form an accusative by suffixing the accusative particle *-ba* (ʾiyarusālem-*bā*).

References

- Bezold, C. 1905. *Kebra Nagast. Die Herrlichkeit der Könige, nach den Handschriften in Berlin, London, Oxford und Paris zum ersten Mal im Äthiopischen Urtext hrsg. und mit dt. Übers. versehen*, Abhandlungen der Philosophisch-Philologischen Klasse der Königlichen Bayerischen Akademie der Wissenschaften, 23/1 (München: Verlag der K. B. Akademie der Wissenschaften in Kommission des G. Franz’schen Verlags, 1905).
- Dillmann, C.F.A. 1899. *Grammatik der äthiopischen Sprache*, 2nd edn, ed. C. Bezold (Leipzig: Tauchnitz, 1899).
- 1907. *Ethiopic Grammar, Second Edition, enlarged and improved*, ed. C. Bezold, tr. J.A. Crichton (London: Williams & Norgate, 1907).
- Tropper, J. 2002. *Altäthiopisch: Grammatik des Ge‘ez mit Übungstexten und Glossar*, Elementa linguarum Orientis, 2 (Münster: Ugarit-Verlag, 2002).
- Zeldes, A. and C.T. Schroeder 2016. ‘*Scriptorium* Part-of-Speech Tagsets for Sahidic Coptic’, Version: 1.1.7_2016.03.14, <http://copticcriptorium.org/download/tools/scriptorium_tagset_documentation.pdf> (last accessed 2 November 2016).
- Zuurmond, R. 2001. *Novum Testamentum Aethiopice, III: The Gospel of Matthew*, Aethiopistische Forschungen, 55 (Wiesbaden: Harrassowitz, 2001).