Hideyuki Inui
Yamaguchi University

Nominal suffixes as markers of information structure in Basketo

Abstract

This paper deals with the information function of two nominal suffixes, *-i* appearing in all nouns, and *-n-* in first- and second-person pronouns in Basketo, a North Omotic language predominantly spoken in the Basketo Special Woreda in Ethiopia. The suffix *-i* is often described as nominative. However, object nouns without definite marker can be marked by *-i*, and as a result *-i* can appear in both subject and object in the same sentence. We analyze morpheme *-i* as a marker of specificity. Suffix *-n-* distinguishes short and long forms of the first- and second-person subject pronoun. The short form is the same as the possessive. In general, possessive does not bear any pragmatic information in discourse. Likewise, short pronouns also show no pragmatic function, but show what is subject or agent in a clause. On the other hand, long pronouns are morphologically and pragmatically marked. We analyze morpheme *-n-* as the foregrounded topic in discourse in contrast with zero anaphora.

Keywords: nominal suffix, personal pronoun, specificity, topic, discourse analysis

1. Introduction

This paper[1] deals with the information function of two morphemes, *-i* appearing on all nouns, and *-n-* in first- and second-person subject pronouns in Basketo[2], a North Omotic language predominantly spoken in the Basketo Special Woreda in Ethiopia. According to the Federal Democratic Republic of Ethiopia Population Census Commission, the number of native speakers of Basketo is estimated at 78,284 (2007 census). Basketo is one of the least studied languages of Ethiopia. There are some recent studies of the language, which deal mostly with morphology. Amha (1993, 1995) deals with noun morphology including personal pronouns. Schütz (2006) analyzes nominative and accusative marking of common and personal pronouns in the framework of Distributed Morphology. Sottile (2002: 90-105), which is a descriptive grammar of Basketo, deals with personal pronouns. Treis (2014) analyzes the grammatical means of encoding interrogativity in Basketo, based on a corpus of recorded spontaneous speech events. However, none of these deal with the information function of *-i* or long and short forms of first- and second-person pronouns discussed here.

2. Methodological preliminaries

The analysis of Basketo adopts the framework of information structure presented in Lambrecht (1994). According to the proposed theory, the most important categories of information structure are: 1) presupposition and assertion, 2) identifiability and activation, 3) topic and focus.

Presupposition and assertion have to do with the structuring of propositions into portions which a speaker assumes an addressee already knows or does not yet know. According to Lambrecht (1994: 51ff.), pragmatic presupposition is the set of propositions lexicogrammatically evoked in a sentence which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered. On the other hand, the pragmatic assertion is the proposition

---

[2]   ISO 639-3 code: bst. Basketo has 29 consonants and 10 vowels as follows: p, t, ʦ, ʧ, k, ʔ, b, d, ʣ, g, p', ʦ', ʧ', k', ɓ, ɗ, ɸ, s, ʃ, h, z, ʒ, ħ, m, n, l, r, w, j, i, e, a, o, u, ii, ee, aa, oo, uu. Acute accent represents high tone.

expressed by a sentence which the hearer is expected to know or take for granted as a result of hearing the sentence uttered. For example, in using a restrictive relative clause as in *I finally met the woman who moved in downstairs*[3], the proposition expressed by the relative clause becomes part of the pragmatic presupposition, namely "old information". The main clause bears a (pragmatic) assertion, namely "new information"[4].

The second category concerns referents. Identifiablity and activation have to do with a speaker's assumptions about the statuses of the mental representations of discourse referents in the addressee's mind at the time of an utterance. According to Lambrecht (1994: 77ff.), identifiable referent is one for which a shared representation already exists in the speaker's and the hearer's mind at the time of utterance, while an unidentifiable referent is one for which a representation exists only in the speaker's mind. Identifiablity has to do with the grammatical categories of definiteness and specificity. Definiteness is a formal feature associated with nominal expressions which signals whether or not the referent of a phrase is assumed by the speaker to be identifiable to the hearer. Specificity has to do with the referent of indefinite noun phrases. A specific indefinite NP is one whose referent is identifiable to the speaker but not to the hearer, while a non-specific indefinite NP is one whose referent neither the speaker nor the hearer can identify at the time of utterance. We will use this framework to discuss the nominal suffix *-i* in subsection 5.1.

On the other hand, activation has to do with consciousness. According to Lambrecht (1994: 93ff.), the psychological factors determining the activation states[5] of discourse referents are thus consciousness and the difference between

---

[3] According to Lambrecht (1994: 55-56), the pragmatic presuppositions lexicogrammatically evoked with the utterance in this sentence can be loosely stated as the following set of propositions: 1) the addressee can identify the female individual designated by the definite noun phrase (by a grammatical morpheme, the definite article *the*), 2) someone moved in downstairs from the speaker (by a grammatical construction, the relative clause *who moved in downstairs*), 3) one would have expected the speaker to have met that individual at some earlier point in time (by a lexical item, the adverb *finally*).

[4] Lambrecht restricts the use of the terms "old information" and "new information" to aspects of information associated with proposition here.

[5] Chafe (1987: 25ff.) defines three different activation states. An active concept ("given information") is one that is currently lit up, a concept in a person's focus of consciousness. A semi-active concept ("accessible information") is one that is in a person's peripheral consciousness, a concept of which a person has a background awareness, but which is not being directly focused on. An inactive concept ("new information") is one that is currently in a person's long-term memory, neither focally nor peripherally active. Lambrecht (1994: 94) refers to what Chafe calls "concept" as "(mental representation of) referent".

short-term memory and long-term memory. An item is active if it is "currently lit up" in our consciousness, and activation normally ceases as soon as some other item is lit up instead. The active state of a referent is formally expressed typically via pronominal coding of the corresponding linguistic expression. The pronominal coding applies to free and bound pronouns, inflectional affixes, and null instantiation (zero coding) of an argument.

The final category concerns relations. Topic and focus have to do with a speaker's assessment of the relative predictability vs. unpredictability of the relations between propositions and their elements in a given discourse situation. Topic is the predictable element in an utterance. Therefore, topic is included in the pragmatic presupposition without being identical to it. On the other hand, focus is that portion of a proposition which cannot be taken for granted at the time of speech. It is the unpredictable or pragmatically non-recoverable element in an utterance. The focus of a sentence is generally seen as an element of information which is added to the pragmatic presupposition. Therefore, focus is part of an assertion without coinciding with it.

Topic referents have a degree of pragmatic accessibility. For postulating a general correlation between the activation and identifiability states of topic referents and the pragmatic acceptability of sentences, we can adopt Givón's scale for the coding of topic accessibility (Fig. 1). The phonological scale runs from zero anaphora to stressed/independent pronouns, and the word-order scale from R(ight) dislocated DEF-NP's to L(eft) dislocated DEF-NP's. As for continuity or accessibility, the left-most element codes more continuous topics, while the right-most less continuous ones[6]. Both Y-movement[7] (contrastive topicalization) and cleft-focus can be considered instances of more discontinuous/surprising topic constructions where the topic is placed to the left of the comment. We will mainly use the phonological scale to discuss nominal suffix *-n-* in subsection 5.2. Zero anaphora is most obvious and picks up the most continuous and accessible topic for the speaker and hearer. We will regard it as a backgrounded topic, in contrast with independent long personal pronouns in Basketo as a foregrounded

---

[6]   According to Givón (1983: 19ff.), the topic-comment orders show higher average numeral values for referential distance than comment-topic orders in some data of several languages.

[7]   The term Y-movement (or: Yiddish movement) was used to describe the fronting of a noun phrase, which was felt to be reminiscent of Yiddish-influenced American English. Givón defines Y-movement as an object-topicalizing construction, where the more topical patient/object is fronted and the less topical agent/subject is postponed (yielding OSV in SVO languages). The question will be further illustrated in examples (30c) and (31c). See also Pekarek, De Stefani & Horlacher (2015: 51).

topic for emphasis. The short personal pronouns are in neutral position relative to backgrounded or foregrounded topic.

Particular elements of this scale are positioned as follows:

most continuous / accessible topic
- zero anaphora
- unstressed / bound pronouns or grammatical agreement
- stressed / independent pronouns
- R-dislocated DEF-NP's
- neutral-ordered DEF-NP's
- L-dislocated DEF-NP's
- Y-moved NP's ("contrastive topicalization")
- cleft / focus constructions
- referential indefinite NP's
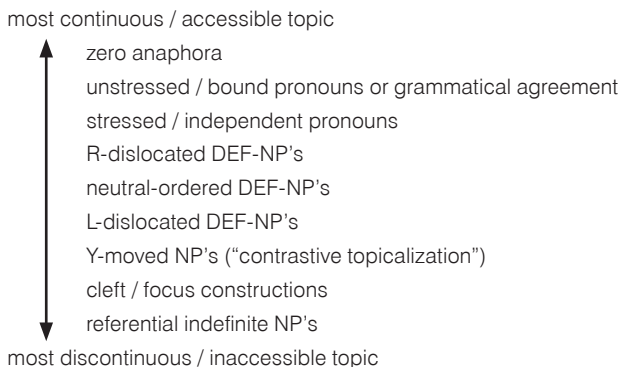
most discontinuous / inaccessible topic

Fig. 1. Scales in the coding of topic accessibility (Givón 1983: 17)

In addition, we will define the linguistic behaviour of the possessive forms of personal pronouns here. Herslund & Baron (2001: 2-3) defines "possession as the linguistic expression of the relation between two entities, a Possessor and a Possessum, such that one, the Possessor, is seen as being in some way related to the other, the Possessum, as having it near or controlling it". Possession can be classified into two main types of linguistic constructions: attributive (e.g. *my credit card*) and predicative (e.g. *I have a credit card*). The former attribute possession is highly general in meaning while the latter predicate possession is more specific[8]. For example, compare **I come home** with *they always prepared for* **my** *coming home*, the former's subject **I** bears the topic automatically in a sentence[9], and the latter's possessive **my** modifies the verbal noun only and does not bear

---

[8]   According to Heine (1997: 28-29), a phrase like *my house* may derive from a large number of underlying sentences such as *I own the house*, *I live in the house*, *I rented the house*, *I built the house*, etc.

[9]   The notion of topic or theme as the first element in the sentence is extensively discussed in Prague School research (Functional Sentence Perspective). See Firbas (1979) about the idea of the communicative dynamism (CD). However, from a viewpoint of typological studies, sentence-initial element can be not only topics but foci from the strategy of left--movement (pre-verbal ordering) of V-initial languages. In any case, sentence-initial element has some pragmatic functions in a sentence.

any pragmatic (namely, topic or focus) information in a sentence unless the possessive has strong (marked) accent in English. Therefore, the possessive forms of personal pronouns discussed here show only grammatical or semantic function at phrase level, and bear no pragmatic information in discourse. We will use this idea for explaining the function of short pronouns of Basketo later.

3. Grammatical outline of Basketo

This section gives a grammatical outline of Basketo focusing on aspects which are relevant for the topic under discussion.

3.1. Terminal vowels

Terminal vowels are the stem-final vowels found in citation forms (called "absolutive forms") of nominals in North Omotic languages generally. In Basketo the terminal vowels are /a/ for masculine nouns[10] and /o/ for the rare feminine nouns[11]. According to Hayward (2001), Basketo has both unstable terminal vowel (UTV) type and stable terminal vowel (STV) type. The former is found in core cases; nominative, accusative and genitive, and the latter in oblique cases; dative, ablative, instrumental etc., and with definiteness. Table 1 shows UTV-type and STV-type of Basketo.

TABLE 1. UTV-type and STV-type of Basketo

as-á 'man' (/ *kaná* 'dog') (ABS)	
UTV-type	STV-type
as-í (NOM)	*asá-bo* 'for a man' (DAT)
as-í (ACC)	*asá-ppo* 'from a man' (ABL)
as-í / kaná (GEN)[12]	*asá-bara* 'by a man' (INSTR)
	asá-da 'the man' (DEF)

---

[10] Terminal vowels immediately following alveolar affricates or fricatives are sometimes hard to hear in spontaneous speech, e.g. absolutive and nominative forms are *ʃóóʃ-a / ʃóóʃ-i* 'snake', *tʃʼuutʃ-á / tʃʼuutʃ-í* 'louse'. However, terminal vowels must drop in the case of non-specific or generic meaning, e.g. *indí garta ʃóóʃ woɗe*. 'There are snakes in mountains'.
[11] Feminine kinship nouns have the terminal vowel /o/ as follows; *índó* 'mother', *aakkó* 'grandmother', *míʃó* 'sister', *mátʃó* 'wife', etc.
[12] Genitive constructions are shown by word order (GN) without changing the terminal vowels in non-human nouns. On the other hand, the terminal vowels change to /i/ like nominative in human and personal nouns, e.g. *kaná gólsa* 'dog's tail', *buná sawá* 'smell of coffee', *baabí ojdá* 'father's chair', *ts'oosí k'áála* 'God's voice'.

3.2. Case marking

Basketo has both case marking and verb agreement. Both case marking and agreement can express the relation between a main verb and its dependent noun phrases within a clause. In languages in general, morphological marking of grammatical relations may appear in either the head or the dependent member of the constituent, in both or in neither. Nichols (1986) calls these four types of marking head marking, dependent marking, double marking, and neutral marking.

For a verb and its dependent nouns there are the following four possibilities (in all distinguished types, heads are indicated by superscript H, affixal markers by M):

Dependent-marked:

Noun_1 + $^M\text{Case}$ Noun_2 + $^M\text{Case}$ $^H\text{Verb}$

Head-marked:

Noun_1 Noun_2 $^H\text{Verb}$ + $^M\text{AFF}_{N1}$ + $^M\text{AFF}_{N2}$

Double-marked:

Noun_1 + $^M\text{Case}$ Noun_2 + $^M\text{Case}$ $^H\text{Verb}$ + $^M\text{AFF}_{N1}$ + $^M\text{AFF}_{N2}$

Neutral-marked:

Noun_1 Noun_2 $^H\text{Verb}$

Basketo has a nominative-accusative system like other Omotic languages. But morphological marking is partly determined by definiteness, as in North Ometo languages. Indefinite nouns are morphologically marked only in the nominative with suffix *-i*. Basketo has a marked nominative system[13] in this case. On the other hand, definite nouns are morphologically marked for both nominative *-di* and accusative *-dani*. However, contrary to other Ometo languages, object nouns without a definite marker can be marked with the suffix *-i* which seems to be

---

[13] König (2006: 658) distinguished two subtypes among the marked-nominative languages with regard to the morphological markedness of nominative and accusative. Type 1 (the more common one), in which the accusative is the morphologically unmarked form and the nominative the morphologically marked form, and type 2, in which both case forms, nominative and accusative, are morphologically marked. In type 1 of marked-nominative languages, the accusative is morphologically unmarked, functionally unmarked and used in citation. In type 2, the accusative is morphologically marked, functionally unmarked, and used in citation. According to König (2006: 687), there is a concentration of type-2 languages within Highland East Cushitic and North Omotic, though Basketo is classified as an accusative language with South Omotic (Hamar, Dime and Aari) in König (2008: 89ff.).

a marker of specificity[14]. This shows a neutral case-marking system. Therefore, Basketo has a split marked-nominative system. Tables 2 and 3 show nominative / accusative case marking of definiteness and specificity in Basketo. We will discuss the specific function of *-i* in the subsections below[15].

TABLE 2. Nominative case marking in Basketo

	+DEF	−DEF
+SPEC	-di	-i
−SPEC	-da	-a

TABLE 3. Accusative case marking in Basketo

	+DEF	−DEF
+SPEC	-dani	-i
−SPEC	-dana	-ø

As an illustration, examples (1-2) are for non-specific (i.e. generic) nominative and accusative, example (3) for marked nominative, example (4) for both marked with definite marker, and example (5) for both nominative and accusative marked by *-i* as a specific marker.

(1) kan-á   áiʃ   múj-íre
    dog-ABS   meat(-ACC)   eat-IMPF
    'Dogs eat meat.'

(2) bíínn-a   gúún   ejts-íre
    mosquito-ABS malaria(-ACC) cause-IMPF
    'Mosquitos cause malaria.'

(3) kan-í   áiʃ   múj-íne
    dog-NOM   meat(-ACC)   eat-PF
    'A dog ate meat.'

---

[14] The suffix *-i* always seems to bear a high tone in a sentence, irrespective of word accent type. This may have relation to pragmatic function or sentence intonation. We must collect more data for ascertaining the relation between tone and case marking.

[15] Feminine nouns show diminutive or lovable meanings with /-in/, and are used frequently in a colloquial sentence, e.g. *awá* 'sun' / *awín* 'lovable sun'. However, such feminine nouns are necessarily specific. Therefore, we will deal here with the masculine nouns only for discussing the difference of specificity.

(4) kaná-**d-í**   áiʃa-**d-ani**   múj-íne
     dog-DEF-NOM   meat-DEF-ACC   eat-PF
     'The dog ate the meat.'

(5) kan-**í**   baw-**í**   éédɗ-íne
     dog-NOM/ACC   cat-ACC/NOM   catch-PF
     'A dog caught a cat.' or 'A cat caught a dog.'

3.3. Verb agreement

Verb conjugation in Basketo shows both subject agreement and aspect. Subject agreement indicates person, gender, and number. Aspect distinguishes imperfective and perfective. Verb conjugation shows "polyfunctionality", expressing person, gender, and number by one portmanteau morpheme, and is highly syncretic. Thus, in the imperfective the suffix **-áre** is used for 1SG, 2SG and 3SG.F and **-íre** for 3SG.M and all plural. There are two Perfectives; the recent past with **-áde** / **-íde** and past with **-íne**[16]. In the former the suffix **-áde** can be used for 1SG, 2SG and 3SG.F and **-íde** for 3SG.M and all plural. With the latter, the suffix **-íne** can be used for all personal endings of perfective, making the agreement a poor guide to the verb's subject. Therefore, subject nouns, especially independent personal pronouns, will be overt. See example (6) for **-áde** / **-íne** and (7) for **-íde** / **-íne**. In Table 4 the syncretic paradigm of Basketo is compared to the fully differentiated paradigm of Wolaytta.

(6) táán-í / néén-í / íz-á   lúkk-**áde** / **-íne**
     1SG / 2SG / 3SG.F-NOM   go-PF
     'I / You (sg.) / She went.'

(7) íj-í / núún-í / jínt-í / ínt-í   lúkk-**íde** / **-íne**
     3SG.M / 1PL / 2PL / 3PL-NOM   go-PF
     'He / We / You (pl.) / They went.'

In sum, Basketo is formally a double marking language at clause level, but in practice a neutral marking language, because of poor guidance from both case marking and verb agreement, as can be seen from example (5).

---

[16] The difference between the two is not significant for this study. See Inui (2017) for details of the tense-aspect system of Basketo.

TABLE 4. Basketo / Wolaytta person marking on main verb

	Imperfective		Perfective	
	Basketo	**Wolaytta**	**Basketo**	**Wolaytta**
1SG	-áre	-aísi	-áde/-íne	-aási
2SG	-áre	-aása	-áde/-íne	-ádása
3SG.F	-áre	-aúsu	-áde/-íne	-aásu
3SG.M	-íre	-eési	-íde/-íne	-iísi
1PL	-íre	-oósi	-íde/-íne	-ída
2PL	-íre	-eéta	-íde/-íne	-ídéta
3PL	-íre	-oósóna	-íde/-íne	-idósóna

3.4. Personal pronouns

First- and second-person subject pronouns[17] in Basketo have short and long forms. Similar pronoun paradigms are found in Ometo languages, including Wolaytta, Gamo, Gofa, and Dawro, as well as in Bench (see Amha 2012: 471). The short form is the same as the possessive form, which is morphologically the simplest form. The long form shows the morpheme *-i* parallel to nouns. The object can be marked by *-na* for all personal pronouns, and proper and kinship nouns. This morpheme may be the survival of the old accusative marker (see Hayward & Tsuge 1998: 22-26). Tables 5-7 show paradigm of proper nouns, kinship nouns, and personal pronouns (boldface for the old accusative marker), respectively.

TABLE 5. Paradigm of proper nouns

	ABS	NOM	ACC
Male name	Wólk'á	Wólk'-í	Wólk'á-ná / Wólk'á-ní
			Wólk'á-**ná**ná / Wólk'á-**ná**ní
Female name	Wotts'anó	Wotts'an-á	Wotts'aná-ná / Wotts'aná-ní
		Wotts'an-í	Wotts'aná-**ná**ná / Wotts'aná-**ná**ní

---

[17] Object personal pronouns, and proper and kinship nouns in Basketo can also have suffix *-n-* as in Tables 5-7. We will discuss the behaviour of object personal pronouns with the suffix *-n-* as a topic marker in 5.4.

TABLE 6. Paradigm of kinship nouns

	ABS	NOM	ACC
Male kinship	baabá 'father'	baab-í	baabá-ná / baabá-ní baabá-**ná**ná / baabá-**ná**ní
Female kinship	índó 'mother'	índ-á / índ-í	índá-ná / índá-ní índá-**ná**ná / índá-**ná**ní

TABLE 7. Paradigm of personal pronouns

	POSS	NOM	ACC
1SG	tá	táání / tá	tááná / táání, táá**ná**ná / táá**ná**ní
2SG	né	nééní / né	nééná /nééní, néé**ná**ná / néé**ná**ní
3SG.M	í	íjí	íjáná / íjání, íjá**ná**ná /íjá**ná**ní
3SG.F	ízí	ízá / ízí	ízáná / ízání, ízá**ná**ná / ízá**ná**ní
1PL	nú	núúní /nú	núúná / núúni, núú**ná**ná / núú**ná**ní
2PL	jíntí	jíntí	jíntáná / jíntání, jíntá**ná**ná / jíntá**ná**ní
3PL	íntí	íntí	íntáná / íntání, íntá**ná**ná / íntá**ná**ní

3.5. Adjective predicates

Although Basketo is said to have a nominative-accusative system, in adjective predicates we find an interesting marking by specificity and definiteness. Adjective predicates typically express features or properties of their subject nouns. Therefore, both inanimate and animate nouns, if indefinite, use the citation (absolutive) form in **-a**. On the other hand, definite nouns are marked by nominative **-i**. Specific human nouns are also marked by nominative. This shows that morpheme **-i** has to do with specificity and definiteness. See example (8a) and (9a) for non-definite and (8b) and (9b) for definite in inanimate and animal nouns.

(8a) mítts-**a**   ɓarínts-e
      tree-ABS   tall-PRD
      'Trees are tall.'

(8b) míttsa-d-**í**   ɓarínts-e
      tree-DEF-NOM   tall-PRD
      'The tree is tall.'

(9a) ɸar-**á**   ɓéttʃ-e
      horse-ABS   big-PRD
      'Horses are big.'

(9b) ɸará-d-**í**   ɓéttʃ-e
      horse-DEF-NOM   big-PRD
      'The horse is big.'

In example (10) human nouns can be marked by either absolute *-a* or nomina-
tive *-i*, but in example (11) divine nouns are always marked by nominative *–i*,
irrespective of definite marker *-d*. Likewise, a generic subject in intransitive takes
absolutive *-a*, too. See example (12) for generic meaning, and example (13) for
specific in intransitive.

(10) as-**á** / -**í**   ɓéttʃ-e
      man-NOM   big-PRD
      'Men are big.'

(11) ts'oos-**í**   kóʃʃ-e
      god-NOM   good-PRD
      'God is good.'

(12) méh-**a**   tʃ'ing-ári   ħájk'-íre
      animal-ABS   get old-CNV.SS   die-IMPF
      'Animals die when they get old.'

(13) génts-**i**   jédd-íre
      ox-NOM   walk-IMPF
      'The ox walks.'

4. Statistical analysis of corpus

Here we use statistical data from a corpus of spoken Basketo, a conversation
between four children in Balt'a village[18]. This corpus of recorded spontaneous
speech consists of 82 clauses[19]. In language description, when we collect tran-

---

[18] Balt'a village is located about 30 minutes by car from the town of Basketo. The infor-
mants of this corpus are children from 8 to 10 years old. Children in the village use Basketo
in everyday conversation, but since the lingua franca is Amharic, they also use Amharic
to communicate with other communities, especially at school. They know one or two
neighbouring languages imperfectly, too.
[19] The breakdown of that is as follows: 58 main clauses, 15 converbs, 3 adverbial clauses,
4 relative clauses, and 1 noun clause. The full text of the corpus is available online at
https://purl.org/JAEL/CEL/lang/bst/r/001 [24.08.20]. Each sentence is numbered and the
alphabetic characters (A-D) identify the four children (e.g. 1-A).

sitive sentences, we usually include both subject and object. On the other hand, recorded natural conversation foregrounds another aspect of language: the information structure.

4.1. Subject marking

We found three types of anaphoric subject form: zero, short, and long forms. See example (14) for zero anaphora, (15) for short form, and (16) for long form.

(14) **ø**   timirtááL   lúkk-ár-bajá?
      (2SG.NOM)   school   go-IMPF-INTRO
      'Are you really going to school?' (40-B)

(15) ħát   **tá**   lúkk-ánda
      now   1SG.NOM   go-INT
      'I'll go now.' (7-C)

(16) **táání**   lúkk-ákkaje
      1SG.NOM   go-PF.NEG
      'I haven't been.' (2-B)

Table 8 shows the appearance of singular first- and second-person pronouns (there are no plural forms). Singular first- and second-person subject referents account for 84% of the total verbs. It shows that it often happens that either the speaker marked by first-person or the hearer marked by second-person is topicalized in the natural sentences.

TABLE 8. Appearance of subject

	1SG	2SG	3SG	Total
Subject	34	35	13	82
%	41	43	16	100

For first- and second-person subjects, statistical data of null, short, and long forms, are shown in Table 9. Null accounts for about two-thirds of the total. On the other hand, long forms account for less than 10% of the total. Therefore, we realize that the long form may be rather marked in discourse.

TABLE 9. Null, short and long of 1SG / 2SG

	Null	Short	Long	Total
1SG	18	11	5	34
	53%	32%	15%	100%
2SG	26	8	0	34
	76%	24%	0%	100%
Total	44	19	5	68
	65%	28%	7%	100%

With third-person subjects, zero anaphora accounts for less than one-fourth of the total, while specific nouns increase more than 50% in Table 10. Therefore, we realize that specific subject nouns may be unmarked in discourse. Specific nouns show earlier topics resumed and marked for the addressee's identifiability and activation.

TABLE 10. Subject marking

Zero anaphora	DEF	SPEC	INDEF	Total
3	1	7	2	13
23%	8%	54%	15%	100%

4.2. Object marking

The types of object marking are shown in Table 11. In general, indefinite nouns function as new information and become the focus of the sentence. On the other hand, zero anaphora picks up an activated referent and becomes the backgrounded topic in discourse. The transitive clauses are 48 of total 82 clauses in the corpus[20]. Implicit objects (zero anaphora) account for 60% of the total and refer to the most accessible (activated) referent, typically the current topic. On the other hand, specific objects are very rare. Definite objects account for more

---

[20] In this corpus, all objects were inanimate with one exception, the first person pronoun *táánáná* in 16-B.

than one-fourth of the total and show an overt topic with possessive or demonstrative.

TABLE 11. Object marking

Zero anaphora	DEF	SPEC	INDEF	Total
28	13	1	5	47
60%	28%	2%	11%	100%

See example (17) for zero anaphora, (18) for specific (*-i*), (19) for definite (*-d*), and (20) for indefinite (*-ø*).

(17) wúú,   né   á   **ø**   erár-dor-a?
      Oh   2SG.NOM   why   (grade)   know-EXCL-INTRO
      'Oh, how do you know (your grade)?' (10-B)

(18) táání   lúkk-á,   **daraʤ-í**   sísk-í-jéj-íne
      1SG.NOM   go-CNV.SS   grade-ACC   hear-CNV.SS-come-PF
      'I went and got my grade.' (3-A)

(19) tá   tá   **daraʤá-da(ni)** hamús   sísk-ánda.
      1SG.NOM   1SG.POSS   grade-DEF.ACC   Thursday   hear-INT
      'I'll get my grade on Thursday.' (57-A)

(20) ats'annaʔ-íno-ppo   ɦaná   **óós**   oots-ár-bajá?
      study-REL.PF-ABL   after   work(-ACC)   do-IMPF-INTRO
      'Will you work after studying?' (50-B)

4.3. NP deletion

Many Omotic languages morphologically distinguish two types of converbs for switch reference: the same-subject converb and the different-subject converb. The former indicates that the converb's subject is the same as the subject of the main verb. The latter indicates that the converb's subject is different from the subject of the main verb[21]. A converb is defined as a nonfinite verb form whose main function is to mark adverbial subordination. In a nonfinite clause the sub-

---

[21] In Basketo there are three converb types: 1) same-subject anterior converb (*-ára* / *-íra*), 2) same-subject simultaneous converb (*-i*), and 3) different-subject anterior converb (*-ín*). We often find the shortened form *-á* from *-ára* for 1SG (lúkk-*ára*) in discourse as in ex (18).

ject cross-linguistically tends to be unexpressed and thus depends for its referential interpretation on the overtly expressed subjects of main clauses. However, we found all four patterns of NP deletion (Table 12) in the data. These data show that NP deletion in Basketo is not a matter of syntax but of pragmatics. We must explain the appearance of subject for examples (21) - (24) from information structure.

TABLE 12. NP deletion

Sub-Clause	Main-Clause	
–	+	(21)
+	–	(22)
+	+	(23)
–	–	(24)

(21) **ø**   jebet sirá   ekk-í,   **tá**   oots-áre
     (1SG.NOM)   homework (-ACC)   take-CNV.SS   1SG.NOM   do-IMPF
     'I'll take homework and do it.' (23-C)

(22) **táání**   lúkk-á,   **ø**   daraʤ-í   sísk-í-jéj-íne
     1SG.NOM   go-CNV.SS   (1SG.NOM)   grade-ACC   hear-CNV.SS-come-PF
     'I went and got my grade.' (3-A)

(23) árt dabtár   **sol-í**   ekk-ín,   **tá**   wong-ákkaje
     art exercise book (-ACC) thief-NOM   take-CNV.DS   1SG.NOM   buy-PF.NEG
     'The thief took the art exercise book, but I haven't bought (a new one) yet.' (63-A)

(24) **ø**   zinááb   lúkk-ín,   **ø**
     (1SG.NOM)   yesterday   go-CNV.DS   (someone)

     sanjí   jéé-bte   géj-íne
     Monday   come-IMPER   say-PF
     'When I went (to school) yesterday, they said to come on Monday.' (59-A)

5. Discussion

5.1. *-i* as a specific marker

Within the Afroasiatic phylum, marked nominative is found in Berber, Cushitic, and Omotic languages. Within Omotic, North Omotic languages show a concentration of marked-nominative languages, especially the Ometo languages

are mostly marked nominative (König 2006: 695-698). Tosco (1994: 236) argues that the nominative *-i* of Basketo has been grammaticalized out of a topic marker. He analyzed the suffix *-i* of Basketo as functioning more like a topic than a nominative marker, unlike that of Wolaytta and Gamo. In the previous section we showed that the suffix *-i* of Basketo can appear on both subject and object in the same sentence. Therefore, we can analyze this morpheme neither as nominative marker nor topic marker. Here we propose analyzing this suffix as a marker of specificity. There is no doubt that *-i* has to do with specificity: the evidence from adjective predicate structure in subsection 3.5 shows this.

Inui (2012) has some examples (25-28) of unmarked object without *-a*. As with *'óós* (work)' of (20), the object does not refer to a specific individual, but implies simply generic or abstract meaning. In such a case, the suffix *-i* is infelicitous.

(25) táání   **áiʃ**   ʃááʃk-íne
     1SG.NOM   meat (-ACC)   grill-PF
     'I grilled some meat.'   (Inui 2012: 178)

(26) táání   ħattábo   misí   **úúɸ**   múj-íne
     1SG.NOM   today   lunch   injera (-ACC)   eat-PF
     'I ate bread for lunch today.'   (Inui 2012: 164)

(27) dejʃ-í   **ɓétɕ**   ɓétɕ-ín,   ʃiiʃ-int-ire
     goat-NOM   dung (-ACC)   defecate-CNV.DS   gather-PASS-IMPF
     'Goats defecate and dung is gathered.'   (Inui 2012: 70)

(28) táání   **jétɕ**   dos-áre
     1SG.NOM   song (-ACC)   like-IMPF
     'I like songs.'   (Inui 2012: 83)

Conversely, the suffix *-i* is obligatory in case of nouns modified by demonstrative, as shown in example (29).

(29) táání   zináábo   ħá   **úúɸ-í**   múj-ára   mađ-íne
     1SG.NOM   yesterday   this   injera-ACC   eat-CNV.SS   become sick-PF
     'I ate this bread yesterday and became sick.'

We can regard the conditions for appearance of the specific marker in question as a hierarchy of individuation or a hierarchy of salience. Salience is not treated as a primitive in itself, but rather as the result of the interaction of a number of factors, such as animacy, specificity, singularity, and concreteness (see Comrie 1989: 199).

In summary, the suffix *-i* suggests that the reference of noun phrase in question is important, relevant for the discourse as a whole. The nouns, either subject or object, are activated by adding *-i* morpheme in discourse. Subject is more frequently marked by this morpheme, because subject is more salient than object in discourse. Here, we propose tentatively that this morpheme functions as specific.

5.2. *-n-* as a topic marker

Here we discuss the information function of the morpheme *-n-* of personal pronouns. It will be useful to utilize Lambrecht's definition of topic and focus (1994) and Givón's scale for the coding of topic accessibility explained in section 2.

The short and long forms of the first- and second-person subject pronouns may be used alternatively in the same context, apparently without any semantic difference in several Ometo languages. So far, no analysis of actual use of the short and long pronouns has been made.

Rapold (2006: 341-363) discussed the various forms of pronoun in Bench, which uses two parameters, long or short, and strong or weak tone[22], giving four combinations. Rapold reports that the long strong pronouns are the most discontinuous subject pronouns: they are typically used to code new or resumed topics, while short strong pronouns signal a higher topic continuity and also code subject focus. On the other hand, short weak pronouns are the most continuous (overt) subject pronouns, and long weak pronouns signal higher topic continuity than long strong or short strong pronouns, but are slightly more discontinuous than the short weak pronouns. The correlation between topic continuity and the various forms of pronouns in Bench is shown in Fig. 2.

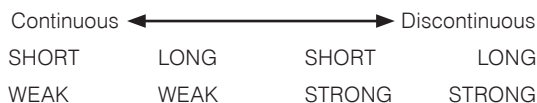| Continuous ◄──────────► Discontinuous			
SHORT	LONG	SHORT	LONG
WEAK	WEAK	STRONG	STRONG

Fig. 2. Topic continuity in Bench

---

[22] Pronouns with the shape CV are termed "short", and those with the shape CVC "long". Pronouns with tone 3 (a neutral mid tone) are termed "weak", those with tone 1 (a salient, extreme low tone), which are pragmatically more marked, "strong".

In this language, focus seems to be determined by stress accent, on the other hand morpheme **-n-** of the Bench long form seems to be characterized as a topic marker. The information structure of four combinations is shown in Table 13.

TABLE 13. The appearance of topic and focus in Bench

	STRONG (+F)	WEAK (−F)
LONG (+T)	T / F	T / Ø
SHORT (−T)	Ø / F	Ø / Ø

T – topic, F – focus

Thinking about what types of nominal are likely to be used as focus and topic, zero marking is used when the referent intended is the most accessible one, generally an activated referent, typically, current topic of conversation. Use of a pronoun guarantees that the referent intended is either activated (especially if unstressed) or at least accessible (if stressed). Use of a definite NP guarantees that the referent intended is identifiable, and generally both inactive and accessible. Use of an indefinite NP generally tells the hearer that the referent is not identifiable in the current context and hence is a new referent being introduced into the context. Thus, zero coding is used for a topic, while realization as an indefinite NP is used for a focal element. Typically, subject has to do with topic, while object has to do with focus. According to Givón (1979: 51-52), in an English text count, 50% of the direct objects were indefinite and 82% of the indefinite NPs were direct objects. We come up with a scale of markedness relations between the form of a referring expression and its function as topic or focus, as shown in Fig. 3.

Markedness of occurrence as focus

←

| Zero | Clitic / bound Pronoun | Pronoun [−stress] | Pronoun [+stress] | Definite NP | Indefinite NP |

→

Markedness of occurrence as topic

Fig. 3. Coding of referents in terms of possible functions
(Van Valin & Lapolla 1997: 205)

Turning to the data of Basketo, zero anaphora, whether of subject or object, functions as a topic, but is backgrounded in discourse. It is important that the short form is the same as the possessive in Basketo. In general, possessive does not bear any pragmatic information in discourse such as explained in section 2. Therefore, short pronouns also have no pragmatic function, but show what is subject or agent in a clause: this form shows only grammatical or semantic function in the clause. On the other hand, long pronouns are morphologically and pragmatically marked. Therefore, morpheme *-n-* makes a foregrounded discourse topic in contrast with zero anaphora, including such topics as normal topic, resumed topic, contrastive topic, unpredictable topic[23], and even focus-like cleft NPs with high animacy. For example, if someone says *It was John that ate the cake*, the referent of the name *John* must already be known to the hearer, namely this is its identifiability status in the mind of the hearer. In this case *someone ate the cake* is the presupposition, *someone = John* is the assertion, the new information, and *John* is the focus of the utterance. Thus, we can consider the continuum from topic to focus as the remit of morpheme *-n-*.

5.3. The use of nominal suffixes *-i* and *-n-* in discourse

Here we will discuss the appearance of both nominal suffixes (*-i* and *-n-*) in the corpus. It may be difficult to find applicable examples from spontaneous speech, but we try to provide evidence for these functions. The following text is the first part of the corpus. We found three types of anaphoric subject form: zero, short, and long forms. Zero anaphora (3-A and 4-B) shows the most activated topic as a backgrounded topic, while long forms (2-B and 3-A) foreground to express the meaning of contrastive topic. On the other hand, short forms (1-A and 4-B) may show the neutral position relative to backgrounded or foregrounded topic.

1-A: **tá**   timirtáál   lúkk-í-jéj-íne
      1SG.NOM   school   go-CNV.SS-come-PF
      'I've been to school.'

2-B: **táání**   lúkk-ákkaje
      1SG.NOM   go-PF.NEG
      'I haven't.'

3-A: **táání**   lúkk-á,   **ø**   daraʤ-í   sísk-í-jéj-íne
      1SG.NOM   go-CNV.SS   (1SG.NOM)   grade-ACC   hear-CNV.SS-come-PF
      'I went and got my grade.'

---

[23] The unpredictable topic nouns are either indefinites entering into the discourse for the first time or definites reentering the discourse after much has intervened.

4-B:	**tá**	giáb	**ø**	sísk-ánda	**ø**	lúkk-ára
	1SG.NOM	tomorrow	(grade-ACC)	hear-INT	(1SG.NOM)	go-CNV.SS

'I'll go and get mine tomorrow.'

Turning to specificity, *-i* marks *daradʒ-í* as specific in 3-A, while zero anaphora is used in 4-B. Likewise, in the following text, the specific suffix *-i* with possessive is used in 8-B, the specific suffix *-i* in 9-C and zero anaphora in 10-B. In both cases, specific nouns[24] may be marked by the suffix *-i*, and if they are activated in discourse, they may become zero anaphora as a backgrounded topic.

8-B:	né	**daradʒ-í**	sínt	késk-ína?
	2SG.POSS	grade-NOM	how much	go up-PF.INTRO

'How much has your grade improved?'

9-C:	**daradʒ-í**	ánd	késk-íne
	grade-NOM	one	go up-PF

'My grade has improved by one.'

10-B:	wúú,	né	á	**ø**	erár-dor-a?
	Oh	2SG.NOM	why	(grade-ACC)	know.IMPF-EXCL-INTRO

'Oh, how do you know (your grade)?'

On the other hand, the specific suffix *-i* does not appear in the non-specific (i.e.generic) meaning as in 50-B.

50-B:	ats'annaʔ-íno-ppo	ɦaná	**óós**	oots-ár-bája?
	study-REL.PF-ABL	after	work(-ACC)	do-IMPF-INTRO

'Will you work after studying?'

51-A:	**ø**	oots-ár
	work(-ACC)	do-IMPF

'Yes, I will.'

5.4. Further evidence from word order

We discussed the information function of two nominal suffixes, and analyzed *-i* as a specific and *-n-* as a topic marker. Finally, we can show some evidence from word order supporting these analyses. The following examples show the simple transitive (a), the corresponding passive voice (b) and the OSV word order (c). In general, the passive construction is a strategy foregrounding the

---

[24] A definite noun such as *daradʒ-ádi* (NOM) / *daradʒ-ádani* (ACC) might be expected because the referent of a phrase is assumed by the speaker to be identifiable to the hearer.

patient, while backgrounding the agent. The patient is promoted from accusative to nominative while the agent is demoted from nominative to oblique case or often deleted. Basketo does not prefer the passive construction but the OSV word order[25]. Moreover, a subject with low animacy (such as 'bedbugs') tends to be avoided. Examples (30c) and (31c) show sentence initial accusative nouns with the suffix *-n-* as a topic marker, on the other hand, the nominative nouns have the suffix *-i* as a specific marker for salience.

(30a) íjí   néénání   bukk-íne
      3SG.M.NOM   2SG.ACC   hit-PF
      'He hit you.'

(30b) ?nééní   í-bara   bukk-int-íne
      2SG.NOM   3SG.M-ABL   hit-PASS-PF
      'You were hit by him.'

(30c) néénáná   íjí   bukk-íne
      2SG.ACC.TOP   3SG.M.NOM   hit-PF
      'You, he hit.'

(31a) ??isí   táánání   ɗak'-íne
      bedbug.NOM   1SG.ACC   bite-PF
      'A bedbug has bitten me.'

(31b) ?táání   isí-bara   ɗak'-int-íne
      1SG.NOM   bedbug-ABL   bite-PASS-PF
      'I has been bitten by a bedbug.'

(31c) táánáná   isí   ɗak'-íne
      1SG.ACC.TOP   bedbug.NOM   bite-PF
      'Me, a bedbug has bitten.'

6. Conclusion

We discussed the information function of two morphemes, nominal suffix *-i*, and *-n-* in first- and second-person pronouns in Basketo. First, though it has been said that Basketo has a nominative-accusative system, the suffix *-i* of Basketo can appear on both subject and object in the same sentence. So, we cannot regard this morpheme as a nominative or topic marker. Here we analyze morpheme *-i* as a specific marker. Second, there are short and long forms of the first- and second-person subject pronouns. The short form is the same as the

---

[25] The same exemplifications are found in Russian. Russian does have a passive construction but its use is less frequent. Russian uses the active with the word order direct object-verb-subject, rather than the passive construction (Comrie 1989: 81).

possessive form. In general, possessive does not bear any pragmatic information in discourse. Likewise, short pronouns also have no pragmatic function, but show what is subject or agent in a clause. On the other hand, long pronouns are not only morphologically but also pragmatically marked. We analyze morpheme ***-n-*** as the foregrounded topic in discourse in contrast with zero anaphora or a short pronoun as the backgrounded topic. For making sure of this idea, it is important to collect more data from natural discourse.

Abbreviations

1	first person	2	second person	3	third person
ABL	ablative	ABS	absolutive	ACC	accusative
CNV	converb	DAT	dative	DEF	definite
DS	different-subject	EXCL	exclamatory	F	feminine
GEN	genitive	IMPER	imperative	IMPF	imperfective
INDEF	indefinite	INSTR	instrumental	INT	intentional
INTRO	interrogative	M	masculine	NEG	negative
NOM	nominative	PASS	passive	PF	perfective
PL	plural	POSS	possessive	PRED	predicate
REL	relativizer	SG	singular	SPEC	specific
SS	same-subject	TOP	topic		

References

Amha, A. 1993. *The case system of Basketo*. Unpublished MA thesis. Addis Ababa University, Ethiopia.

Amha, A. 1995. "Case in Basketo". *African Languages and Cultures* 8. 1-17.

Amha, A. 2012. "Omotic". *The Afroasiatic Languages*, ed. by Z. Frajzyngier & E. Shay. Cambridge: Cambridge University Press. 423-504.

Chafe, W. 1987. "Cognitive constraints on information flow". *Coherence and grounding in discourse*, ed. by R. Tomlin. (Typological Studies in Language 9). Amsterdam: John Benjamins. 21-51.

Comrie, B. 1989. *Language universals and linguistic typology*. Chicago: The University of Chicago Press.

Firbas, J. 1979. "A function view of 'ordo naturalis'". *Brno Studies in English* 13. (Sborník Prací Filozofické Fakulty Brnenské Univerzity 1). 29-59.

Givón, T. 1979. *On understanding grammar*. New York: Academic Press.

Givón, T. 1983. *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins.

Hayward, R. J. 2001. "A further consideration of terminal vowels in Ometo". *New data and new methods in Afroasiatic linguistics: Robert Hetzron in memoriam*, ed. by A. Zaborski. Wiesbaden: Harrassowitz. 53-63.

Hayward, R. J. & Y. Tsuge. 1998. "Concerning case in Omotic". *Afrika und Übersee* 81. 21-38.

Heine, B. 1997. *Possession: Cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press.

Herslund, M. & I. Baron. 2001. "Introduction". *Dimensions of Posssession*, ed. by I. Baron, M. Herslund & F. Sørensen. Amsterdam: John Benjamins. 1-25.

Inui, H. 2012. "Basic verb sentences for Ethiopian language research" [in Japanese]. *Studies in Ethiopian Languages* 1. 48-211.

Inui, H. 2017. "Tense-aspect system in Basketo". *Journal of General Linguistics* 20. 1-35.

König, C. 2006. "Marked nominative in Africa". *Studies in Language* 30(4). 655-732.

König, C. 2008. *Case in Africa*. Oxford: Oxford University Press.

Lambrecht, K. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. New York: Cambridge University Press.

Nichols, J. 1986. "Head-marking and dependent-marking grammar". *Language* 62. 56-119.

Pekarek, D. S., E. De Stefani & A. Horlacher. 2015. *Time and emergence in grammar: Dislocation, topicalization, and hanging topic in French talk-in-interaction*. Amsterdam: John Benjamins.

Rapold, C. J. 2006. *Towards a grammar of Benchnon*. Unpublished PhD thesis. Leiden University, The Netherlands.

Schütz, J. 2006. "Kasusmarkierung im Basketo: Eine Analyse im Rahmen der Distribuierten Morphologie". *Linguistische Arbeitsberichte* 84. 63-75.

Sottile, R. 2002. *Schizzo grammaticale del Basketo (Ethiopia Suboccidentale)*. Unpublished PhD thesis. Università degli Studi di Napoli "L'Orientale", Naples.

Tosco, M. 1994. "On case marking in the Ethiopian language area (with special reference to the subject marking in East Cushitic". *Sem, Cam, Iafet Atti della 7a Giornata di Studi Camito-Semitica e Indeuropei*, ed. by V. Brugnatelli. Milano: Centro studi camito-semitici. 225-244.

Treis, Y. 2014. "Interrogativity in Basketo". *Explorations in Ethiopian Linguistics: Complex predicates, finiteness, and interrogativity*, ed. by R. Meyer, Y. Treis & A. Amha. Wiesbaden: Harrassowitz Verlag. 41-77.

Van Valin, R. D. Jr. & R. J. LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.