The endophoric demonstrative *weeŋ* in Tima – A qualitative corpus analysis

Milo Reinmöller

1 Introduction

The Sudanese language Tima (Niger-Congo) has three demonstratives, one of which does not indicate notions of physical distance: *weeŋ*. In this descriptive work, a corpus of 12 narrative stories is examined in order to describe the functions of this demonstrative. Specifically, on the basis of a qualitative analysis, three functions are proposed, all of which are related to a common semantic-pragmatic profile: the ability of *weeŋ* to signal the low cognitive activation of referents.

Cross-linguistic research has not only shown that all languages studied so far have demonstrative-like linguistic units (Peeters et al. 2021: 411f.), but also that their form and functions are subject to immense variation. Languages differ in the number of demonstrative categories, in the morphosyntactic forms they take (Peeters et al. 2021: 411f.) and in the functions they perform (Himmelmann 1996). Although most research (e.g., Diessel 1999, Peeters et al. 2021) suggests that the most basic function of demonstratives is to point linguistically to objects or persons within the utterance situation, i.e., to perform exophoric reference, Himmelmann's (1996) study in particular has shown that demonstratives perform various functions at the text-internal (endophoric) level as well. There, among other things, they serve information-structural purposes, i.e., they help the listener to identify the right referents (Himmelmann 1996: 226).

Tima, a Niger-Congo language, provides evidence for a demonstrative that does not serve to deictically point to entities present in the utterance situation: the morpheme *weeŋ*. Although paradigmatically related to two demonstrative roots used to linguistically point to proximate and distant referents, this morpheme does not itself convey notions of physical distance, but rather serves to mark referents as known to the listener. The present study investigates the properties of *weeŋ* in order to contribute to the understanding of functions that have often not been perceived as prototypically performed by demonstratives.

To this end, a corpus of 12 narrative stories[1] is examined. Specifically, three different functions of *weeŋ* will be established on the basis of corpus examples. These are the reference to textually distant antecedents (SECTION 2.1.1), the resolution of ambiguous reference where it picks up the least activated of competing referents (SECTION 2.1.2) and the marking of referents known outside the scope of the discourse (SECTION 2.2). In addition, occasional quantitative distributions will be presented to support the hypothesis formulated.

Before proceeding to the main issue of this study, i.e., the presentation of the functions performed by *weeŋ*, a number of introductory sub-sections follow. First, a brief overview of Tima is given in SECTION 1.1; previous research on the three demonstratives in Tima is then presented in SECTION 1.2. In SECTION 1.3, I introduce the theoretical approaches to the analysis of demonstratives that are used in this study. The kind of data and its annotations that serve to investigate *weeŋ* are described in SECTION 1.4. In SECTION 2, the observed functions of the demonstrative *weeŋ* are presented and discussed on the basis of corpus examples. SECTION 3 discusses the findings of this study, with a particular focus on a common pragmatic profile implied by all three functions of *weeŋ* in SECTION 3.2. Finally, a short summary follows in SECTION 4.

1.1 General information on Tima[2]

Tima is a highly endangered language spoken by approximately 7,000 speakers in the Nuba Mountains of Sudan. The number of those Tima who have left their home area remains unclear (Meerpohl 2012: 23). Dimmendaal (see, e.g., 2014: 246; 2018: 383) classifies Tima as one of the Katloid languages, along with the closely related languages Katla and Julut. The Katloid languages are part of the Niger-Congo phylum.

---

[1] All recordings under analysis are stored in the DOBES archive (http://dobes.mpi.nl/) of the Max Planck Institute for Psycholinguistics. These are: 02_AliTia_1; 03_AliTia_2; 08_Hamad_1; 09_Hamad_2; 10_Hamad_311_Hamad_4; 010207_Jenge_LionHyena; 011007_11_AdlaanMisiria_Myth; 011007_14_AdlaanMisiria_Horsequarrel; 031007_Daldum_ClanDividing; 280117_10_Hamid_Clandividing; and 20190108_HamidPearFilm.

[2] Numerous publications on Tima provide an overview of phonetics/phonology and morpho-syntactic properties of the language. The interested reader is referred to Dimmendaal (2009), Bashir (2010) and Alamin (2012), followed by more specific publications. For a social-anthropological study see Meerpohl (2012).

Tima makes use of lexical and grammatical tone and has advanced tongue root vowel harmony.[3] The language is attested to have a flexible word order, varying from AVO, OVA, VAO to AOV in transitive clauses, depending on the pragmatic context. Verbs in Tima can show great morphological complexity, as they have a total of 13 slots for bound elements (Dimmendaal 2014: 246f.). Nouns, on the other hand, usually consist of a prefix containing information about number, followed by a nominal root. Nevertheless, nouns can be accompanied by several proclitics, such as directional or instrumental markers, as well as demonstrative enclitics, which will be presented in detail in SECTION 1.2. Furthermore, nouns can be modified by possessive pronouns, adjectives and quantifiers.

Verbal arguments can be realized in different morphosyntactic forms in Tima, first of all as independent external noun phrases or as independent pronouns. Subjects and objects of transitive clauses are generally indexed by bound forms on the verb, with third person referents indicated by zero forms. Thus, core arguments are often neither overtly indexed on the verb nor represented as free pronouns or noun phrases. Beneficiaries and instrumental arguments may also be left unmentioned. In that case, the verb is marked for the applicative or the instrumental marker, but the slot for the pronoun or noun phrase remains vacant.

After this brief introduction to the socio-cultural setting and the most basic grammatical structures relevant for the understanding of the examples, we will now turn to the demonstratives of Tima, as presented in previous research.

1.2 Demonstratives in Tima

Tima has three demonstrative morphemes, *nʌ/na*, *yaa* and *weeŋ*, which are differentiated neither for number nor for gender.[4] Two studies, namely Alamin (2012) and Dimmendaal & Schneider-Blum (in preparation), propose basic semantic, pragmatic and information-structural functions of the demonstratives. These functional descriptions form the basis of the present analysis and are therefore briefly summarized in SECTION 1.2.1. In SECTION 1.2.2, the two morphosyntactic forms in which demonstratives are realized are presented, i.e., as clitic attachments to lexical roots and in their status as the roots of

[3] In the Tima orthography, as developed by the Tima Language Committee in collaboration with linguists from Khartoum and Cologne, tone is not written. In accordance with this convention, the materials I worked with were not marked for tone. Since I didn't have sufficient experience with the language to transcribe tone, the examples cited are not marked for tone.

[4] The variation between *=nʌ* and *=na* depends on the ATR property of the noun's root vowel, while the independent root of the proximate demonstrative is *nʌ; yaa* and *weeŋ* are invariable.

independent demonstrative pronouns. Thus, the demonstrative morpheme is, in either case, a bound form. Since its status varies between enclitic and bound root, morpheme boundaries (= or -, respectively) are not indicated in the main body of the text when referring to the two functions. Throughout this article, the terms *demonstrative*, *demonstrative form* and *demonstrative morpheme* will be used to refer to both morphosyntactic forms, i.e., demonstrative clitics and demonstrative roots.

1.2.1 Functions

As numerous examples show, *nʌ/na* and *yaa* can, among other functions, have exophoric reference. They can refer to either near (*nʌ/na*) or distant (*yaa*) entities that are present in the immediate environment of the speech situation. The third demonstrative, *weeŋ*, on the other hand, cannot be used for exophoric reference, but is described as referring to previously mentioned referents (Dimmendaal & Schneider-Blum, in preparation).

The present study follows Dimmendaal & Schneider-Blum's analysis, where *weeŋ* is classified as a demonstrative even though it does not have exophoric reference. This is because it is paradigmatically related (see SECTION 1.2.2) to the other two demonstratives, *nʌ/na* and *yaa*, which both have distance-indicating, deictic functions. Therefore, I treat *weeŋ* as a demonstrative, following Himmelmann (1996: 211), who notes that "in several languages, there are elements which share highly specific morphosyntactic features with distance-sensitive demonstratives and, for this reason, have to be considered demonstratives, though distance is irrelevant to their semantics".

Apart from their exophoric or anaphoric uses, the three demonstratives most certainly play a role in the marking of definiteness and specificity. As Dimmendaal & Schneider-Blum (in preparation) put it, "the noun that is attached by the demonstrative clitic always refers to a specific, particular referent", whereas nouns with no demonstrative attached remain vague in that respect. Elsewhere, they note that "the demonstrative clitic […] provides the noun with a definite notion". While Dimmendaal & Schneider-Blum relate their observation mainly to nouns marked with the proximal demonstrative, the present corpus confirms that referents realized with a word containing the non-exophoric demonstrative *weeŋ* are also specific and particular. Tima has no specific definite articles, but the use of demonstratives generally implicates the notion of definiteness. Note, though, that in several of the examples below which contain complex noun phrases, the head noun is encliticized by the demonstrative while selective marking occurs on the modifier of the noun phrase (see (1), (3) and (6)). In such cases, definiteness is not necessarily as clear as it

is here (see, e.g., example (47) in Becker & Schneider-Blum 2020).[5] How demonstratives and the selective marker interact, and whether they express definiteness or rather specificity (see von Heusinger 2002), remains to be investigated.

1.2.2 Morphosyntactic forms

The demonstrative morphemes *nʌ/na*, *yaa* and *weeŋ* are attached to many different types of morphemes and thus belong to a variety of different parts of speech. However, their embeddings can generally be grouped as follows. On the one hand, they attach themselves enclitically to nominal, adjectival or numeral hosts, and on the other hand, they form the roots of independent demonstrative pronouns (see TABLE 1). In this section, the structural properties of both embeddings are briefly described.

PRONOMINAL CLITIC	SEMANTIC BASE	SINGULAR PRONOUN	PLURAL PRONOUN
=nʌ/=na	*nʌ/na*	*cí-↓nʌ́/cííŋ*	*í-↓nʌ́/ííŋ*
=yaa	*yaa*	*cí-yáà*	*í-yàá*
=weeŋ	*weeŋ*	*cú-↓wééŋ/kú-↓wééŋ*	*í-↓wééŋ*

TABLE 1: Tima demonstrative pronominals

Free demonstrative pronouns in Tima consist of one of the three demonstrative roots (i.e., the semantic base) combined with a number-differentiating prefix. All three categories are attested as able to be combined with either a singular or a plural affix. For the proximate demonstrative, a second form is attested, i.e.,

[5] Becker & Schneider-Blum (2020: 27, example (47)):

kùlʌ́, *ìhwáá=ˈná* *ì háˈwúk=í =yɛ́* *ù=ṭɔ̀ndɔ̀ɔ̀;*
yesterday people=DEM.PROX many=SEL=FOC.PL DIR=road

ínʌ̀ *ìháˈwúk=í* *ú-kɔ̀kwɛ́ɛ́* *ídék,* *ì bɛ̀ ʔɛ́ ŋ=í*
PL:DEM.PROX many=SEL 3-hold:PST necks few=SEL

í-pʌ̀k-ʌ̀k-ʌ̀ʌ̀ *yʌ́wùh*
3-throw:PST-AP-INS stones
'Yesterday, lots of people were in the street; (while) most of them were peaceful, *some* threw stones.'

cííŋ in the singular and *ííŋ* in the plural. TABLE 1 shows the full paradigm of demonstrative pronominals (with the morpheme boundaries indicated).

Apart from their presence within independent pronouns, the three demonstrative morphemes are attached as enclitics to other nouns or noun-modifying elements, i.e., adjectives, numerals and nominal modifiers, as well as to phrasal verbal modifiers.[6] In (1), for example, the demonstrative *na* encliticizes to the noun *t̪amaa* 'talk'.

(1) *ʊ-dɔɔ-w-aa* *ɲ=ɪrba* *ə-dah=ɪɪ* *ii-murik*
P-stand.up-EP-INS ERG=Irba P-say=APP APP:PL-Tima

mɛ=yɛ *ɨ-ri-y-aa=t̪aŋ* ***t̪-amaa=na***
OPT=REP P-change-EP-INS=LOC3P SG-talk=DEM1

du-murik=i *twar=a=t̪aŋ* *a=t̪-amaa=na*
MOD-Tima=SEL different=SOUR=LOC3P SOUR=SG-talk=DEM1

də-maadaŋ=ɪ […]
MOD-Katla=SEL
'Then Irba told the Tima people to change the Tima language different from the Katla language […].' (280117_10_Hamid_Clandividing 073-075)

The referents marked by demonstrative clitics (as well as by demonstrative pronouns) seem not to be semantically restricted. They are attached to physical as well as non-physical referents, animate as well as inanimate entities and human as well as non-human referents, as shown in (1)-(4).

(2) *[…] u-kumun-aa* *caak-aa=t̪aŋ=ɪɪ* ***ɲ=ɪhr=na***
P-find-INS become-INS=LOC3P=APP ERG=milk=DEM1

ɲ=i-kʌk
INS=PL-bitter
'[…] and found out that the milk had become bitter (for them).' (280117_10_Hamid_Clandividing 047)

(3) ***ku-juur=nʌ*** *i=i-murik=i* *ɪrba*
SG-magician=DEM1 DIR=PL-Tima=SEL Irba

[6] For a demonstrative with a phrasal verbal modifier, see, e.g., example (8) in Schneider-Blum (this volume).

dʌŋ=ʌ *k-ʌhu*
like.this=FOC SG-name
'The Tima magician is called Irba.' (280117_10_Hamid_Clandividing 042)

In terms of their syntactic function, nominal phrases marked by a demonstrative clitic are also variable. They can perform core functions, such as the role of subject (as in (2) and (3)) or object (as in (1)), as well as oblique functions. In (4), for example, the clitic *=na* attaches to the source-marked head noun of the complex noun phrase *ayamaana ɨkʌliyʌ* 'from the *right* way of speaking'.

(4) ***a=y-amaa=na*** *ɨ=k-ʌli=yʌ*
SOUR=PL-talk=DEM1 DIR=SG-truth=FOC
'from the *right* way of speaking' (07_MusaBukur 001)

Demonstrative clitics are consistently attached to the head noun in complex noun phrases, as in (3) and (4). In addition, they can be attached to both the head noun and the modifying adjective, as in (5). In the case of double marking, the referents are visible/present, i.e., the meeting is still going on (Schneider-Blum p.c.).

(5) *ɪhwaa=na* ***ɪhɪɪk-ə=na***
people=DEM1 two-EP=DEM1

an-tikihiṭ-ʌk *idʌ*
3PRF-arrange.secret.meeting:PLUR-AP bodies
'These two people have arranged to meet secretly (now they are sitting together).' (07.04.09, 2_01-04.wav)

The three demonstrative clitics can also be attached to modifying nouns within complex noun phrases, as in (6). Note that in these cases the demonstrative clitics of the head noun (*ibʌweeŋ*) and the modifying noun (*iiwʌwuŋnʌ*) may differ. This is possible because the demonstrative clitics of modifying nouns do not specify the referent of the head noun, but rather the noun they attach to, i.e., *=nʌ* modifies the noun *iwʌwuŋ* 'grandchildren' in (6).

(6) *ihʌhunen=e* *piir* ***i-bʌ=weeŋ*** *kʌwun*
women=FOC.PL dance PL-child=DEM3 of.course

i=i-wʌwuŋ=nʌ *ɨ=pɨnʌ-y=i*
DIR=PL-grandchild=DEM1 DIR=PRON3SG-EP=SEL
'The women dance, those children, of course, of her grandchildren.' (03_AliTia_2 057)

Finally, the three clitics are also attested as attaching to personal pronouns. These typically include first and second person independent forms.[7] In (7), for example, *=nʌ/=na* is attached to a first person plural inclusive pronoun. According to a native speaker, the clitic emphasizes that the speaker is referring to the community members present in the immediate surroundings of the utterance situation. A realization of the pronoun without the demonstrative clitic would be possible, but would imply that the first person pronominal form refers to all Tima speakers, i.e., also to those who are not present in the utterance situation.[8]

(7) ***mɛð=na*** *a=tintiiliŋ=ʌ*
1PL.INCL=DEM1 SOUR=Tintiiling=FOC

i-tulu-uŋ=nɛð
PL-leave.together-VENT=1PL.INCL
'This us, from Tintiiling, we came out (i.e., we all came out from Tintiiling).' (03_AliTia_2 067)

A final note should be added. Whereas all morphosyntactic contexts presented so far are productive, the three demonstrative roots are also part of several lexicalized temporal or spatial adverbs, such as *aduweeŋ* 'since', exemplified in (8).[9] These constructions will not be discussed in the present analysis.

(8) ***aduweeŋ*** *i-di-y-ʌŋ-aa* *ɲ=ihinʌ*
since P-walk-EP-VENT-INS ERG=PRON3PL
'Since (that time) they came here.' (09_Hamad_2 001/002)

1.3 The analysis of demonstratives

Previous research has established a wide variety of theoretical constructs for analysing demonstratives. This section briefly addresses which of these will be used to describe the functions of *weeŋ* in the following sections.

A classification already used above is the distinction between exophoric and endophoric reference. While the referential domain of exophoric demonstratives is within the utterance situation, the referential domain of endophoric demonstratives is within a discourse (Finkbeiner 2018: 192). This distinction is valuable for the analysis of demonstratives in Tima, as it serves to differentiate *weeŋ* from the other two demonstratives. While *nʌ*/*na* and *yaa* can have

[7] Note that the third person singular and plural pronouns are supposed to be connected to demonstrative pronouns, as argued by Schneider-Blum (2013b: 290f.).
[8] HKD_20230122_metalinguistic-comment_01
[9] The other two are *adɔna* (with its variant *adaana*) 'since' and *adıyaa* 'ever since' (Schneider-Blum 2013a: 28).

exophoric reference, *weeŋ* cannot. We can thus narrow down the range of theoretical approaches to those that deal with the analysis of endophoric demonstratives to investigate *weeŋ*.

The functions of endophoric demonstratives have been addressed by a large number of studies from different perspectives. While some research has looked at the use of endophoric demonstratives for structuring discourse (Cornish 2018; Næss et al. 2020), others have focused on the ability of demonstratives to help addressees to identify the correct referents, thereby describing information-structural properties (Ariel 1990; Gundel et al. 1993). Information-structural approaches are well suited to explaining two uses of *weeŋ* (see SECTION 2.1). In particular, several variables serving to describe the information-structural status of referents introduced by Ariel (1990) will be presented in SECTION 2.1.2, as they help to determine the function of *weeŋ* for disambiguating references.

In addition, Himmelmann's (1996) taxonomy of demonstrative uses in narrative corpora will serve as an important basis for analysis. This study proves to be valuable as it examines linguistic data similar to those in the present investigation, i.e., recordings of monolingual narratives (Himmelmann 1996: 207), and thus identifies several functions that deviate from the often described exophoric reference of demonstratives. Specifically, Himmelmann (1996) conducts a qualitative analysis of demonstrative functions in narratives of five languages, resulting in the description of four potentially universal uses: the *situational*, the *discourse deictic*, the *tracking* and the *recognitional use* of demonstratives (Himmelmann 1996: 240).[10] While the *situational use* describes the function of demonstratives to point exophorically to entities in the immediate surroundings of the narrative (Himmelmann 1996: 219-224), the *discourse deictic use* outlines the property of some demonstratives to summarize previously mentioned events or propositions into a single linguistic unit, thereby creating new discourse referents (Himmelmann 1996: 224ff.).[11] The other two potentially universal uses of demonstratives, i.e., *tracking* and *recognitional use*, will be presented in more detail later, where they will be shown to be applicable to several functions performed by *weeŋ*.

Finally, Peeters and colleagues' (2021) more recent approach to the analysis of endophoric demonstratives deserves a mention. By proposing a top-down model of the factors that influence the choice of demonstrative category in discourse,

---

[10] Himmelmann (1996: 207ff.) studied narratives of English, Ik (Kuliak, Uganda), Nunggubuyu (non-Pama-Nyungan, Northern Australia) Tagalog (Austronesian, Philippines) and Indonesian (Austronesian, Indonesia).

[11] The situational use includes what has been labelled as *Deixis am Phantasma* by Bühler (1934: 121ff.): the notion of the physical distance of a referent from the perspective of a fictitious protagonist (Himmelmann 1996: 222).

Peeters et al. (2021) assign importance to a function that has rarely been considered: the use of demonstratives as a reflection of the interactional space of speaker, topic and addressee. According to the authors, speakers use demonstratives to psychologically move a topic within the interactional space of interlocutors. While proximal demonstratives reflect a close psychological space between topic and speaker and a greater distance between topic and addressee, distal demonstratives psychologically move the topic away from the speaker toward the addressee (Peeters et al. 2021: 422). As a result, this model is able to account for quantitative differences in demonstrative distribution across discourse genres: while some genres, such as scientific articles, use more proximal demonstratives to reflect a greater psychological distance between topic and addressee, other genres, such as narratives, use more distal demonstratives to reflect a close psychological distance between topic and addressee, thereby encouraging interaction by the addressee (Peeters et al. 2021: 420-424).

Interactional space may also influence the choice of demonstrative form in Tima. It is apparent that the texts in this corpus vary considerably in the number and categories of demonstratives used. For example, the retelling of the Pear Story (145 words) contains seven instances of *weeŋ*, while other narrative stories, e.g., the monologue *011007_11_AdlaanMisiria_Myth* (223 words), with an equal number of protagonists, contain no instances of *weeŋ* at all. Thus, it can be hypothesized that different communicative settings contribute to these quantitative differences.

This study focuses on the information-structural properties of *weeŋ*. However, it will be useful as a next step to critically reflect on the results of this study in follow-up research, focusing on quantitative differences regarding the distribution of demonstratives in different genres and settings. As a result, it will be possible to investigate whether speakers of Tima use different demonstrative categories to move referents in their interactional space, as postulated by Peeters et al. (2021).

1.4 Data and method

To identify the functions of *weeŋ*, a corpus consisting of 12 monologues by six mother-tongue speakers of Tima was examined. All narratives were gathered during fieldwork between 2007 and 2019 by a team of linguists from Khartoum and Cologne. These monologues are narrative stories, all involving multiple protagonists, and can be grouped into the following types: stories about anthropomorphic animal characters, mythical stories about the origins of the Tima people, stories involving local community members and a retelling of the Pear Story (Chafe 1980). Audio and ELAN (Max Planck Institute for Psycholinguistics 2022) files were available for analysis. The ELAN files

contained transcriptions, glosses, translations, GRAID annotations (Haig & Schnell 2014) and RefIND annotations (Schiborr et al. 2018).[12] On the basis of these files, additional annotations were performed to provide quantitative evidence. Specifically, all words containing *weeŋ*, i.e., demonstrative pronouns or noun phrases, were automatically extracted and integrated into a spreadsheet containing a total of 54 tokens of *weeŋ*.[13] Of these, four are roots of independent demonstrative pronouns and 50 are clitics attached to other hosts. Additional information was manually annotated within the spreadsheet. The variables and their variants are shown in TABLE 4 in the appendix. Finally, this spreadsheet was imported into RStudio (RStudio Team 2020), where the R programming language (R Core Team 2023) was used to compute and graph the distributions, as discussed in the following sections.

On the basis of this data, the functions of *weeŋ* are identified and described with individual corpus examples. In addition, information and elicitations from a mother-tongue speaker of Tima are used to test the formulated hypothesis.[14] Finally, quantitative evidence is occasionally provided to test or emphasize the assumptions made on the basis of the qualitative analysis.

2 Qualitative analysis

In the following sections, the functions of *weeŋ* will be described. While, in SECTION 2.1, two uses will be introduced that relate to anaphoric distance, the ability of *weeŋ* to refer to entities known outside the discourse will be shown in SECTION 2.2.

2.1 Functions related to anaphoric reference

The two functions presented in this section have in common that the referents marked by *weeŋ* have been mentioned in the previous discourse. First, the function of *weeŋ* in marking long anaphoric distances will be elaborated (SECTION 1.2.1), followed by a description of its use in resolving ambiguous references, where it is argued to select the textually more distant of several

[12] GRAID annotations mainly provide information about the syntactic functions of constituents, the basic semantic profile of referents and the morphological structures of words. RefIND annotations index discourse referents. These indexes allow the tracking of referents across narratives.

[13] Lexicalized adverbs including demonstrative roots (see SECTION 1.2.2) have been excluded.

[14] At this point, I would like to thank Hamid (HKD), a mother-tongue speaker of Tima, whose information and elicitations were extremely valuable for this study. I would also like to thank Gertrud Schneider-Blum, who provided the linguistic data and metatextual information used in this analysis and shared her evaluation of several issues discussed in this paper with me.

potentially denoted referents (SECTION 1.2.2). Throughout this analysis, individual instances of *weeŋ* are assigned to only one of these functions. However, the two functions may overlap, since in some cases *weeŋ* is attached to a referent that, taken individually, has a long distant antecedent and can potentially refer to multiple referents simultaneously. In these cases, instances were assigned to referential disambiguation when the semantic-pragmatic context of the sentence in question allowed for several competing referents to be referred to.

2.1.1 Referring to long distant antecedents

The morpheme *weeŋ* is often attached to anaphoric elements that have long distance antecedents. The textual distance of all instances where referents marked with *weeŋ* have textual antecedent averages 11.52 clauses. FIGURE 1 displays the distribution of the measured textual distances in more detail.

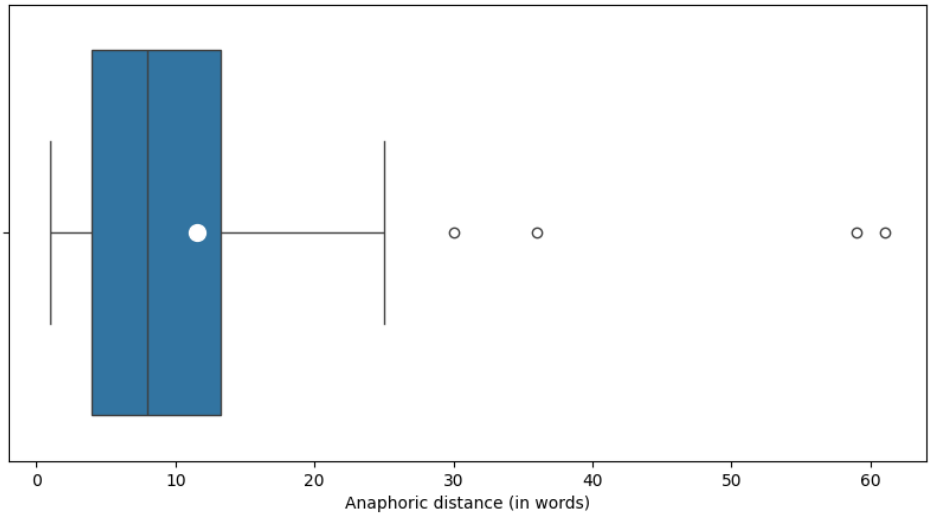

FIGURE 1: Textual distances (in words) from anaphoric forms including *weeŋ* to their antecedents

An example of a long-distance anaphora marked by =*weeŋ* is given in (9), where the nominal phrase *ɪhaamweeŋ* 'that honey' takes up the referent *ɪhaam* 'honey' that was last mentioned 36 clauses ago.

(9) a. *ɪ-cɪ* *ʈuh-uŋ* ***ɪ-haam***
P-go pull-VENT PL-honey
'They went to get honey.' (08_Hamad_1 010)
[…]

b. *ɪ-cɪ* *hʊndɔnɔ-w-aa* ***ɪ-haam = weeŋ***
P-go sit.down-EP-INS PL-honey = DEM3

i̱-kʌl-uk = a = i̱taŋ *ɲ = ihinʌ*
P-chew-CAUS = SOUR = LOC3P ERG = PRON3PL
'They went to sit down with that honey and they ate it.'
(08_Hamad_1 048)

Note that, in this example, *ɪhaamweeŋ* 'that honey' has no competing referents, i.e., there are no other referents of the same semantic category (honey) introduced into the discourse universe. Therefore, it can be excluded that = *weeŋ* in this case serves to help the listener to choose between several possible referents of the same kind. Rather, it seems to remind the addressee that the denoted entity should be known to the hearer, as it was mentioned in the previous discourse. In other words, the speaker seems to have added = *weeŋ* in order to make sure that the listener connects the anaphora with its antecedent and thus does not consider *ɪhaam* 'honey' as a newly introduced referent. In this way, *weeŋ* performs an essential function in creating coherence in the narratives under study: it indicates the coreferentiality of two nominal elements separated by a long anaphoric distance.

Note that the function described above is the most frequently performed function of *weeŋ* in this corpus. A total of 28 tokens (52%) of *weeŋ* have a function similar to the one shown in (9).

2.1.2 Resolving ambiguous reference

The previous section has introduced a function of *weeŋ* in which referents can be described as merely accessible, as they refer to textually distant antecedents. This section will show that referents marked with *weeŋ* may not be individually difficult to recall, but rather less present in the listener's memory than other competing referents. I will argue that, in these cases, *weeŋ* serves to resolve an ambiguous reference by selecting the referent that is more unexpected, or, in other words, less present in the listener's memory.

Before providing evidence for this use of *weeŋ*, it seems necessary to clarify which parameters potentially determine which of the competing referents is less present (or less prominent) in the listener's memory and can thus be used to predict which one of the competing referents is indicated by *weeŋ*. Ariel (1990) discusses the issue of parameters that influence the cognitive accessibility of referents in detail, as I will briefly summarize below.

In her study, Ariel (1990) examines the anaphoric forms used in Hebrew and English texts, focusing on the former. Essentially, she identifies four factors that influence the accessibility of discourse referents: *distance*, *saliency*, *unity* and

number of competing referents (Ariel 1990: 28f.). The parameter *distance* distinguishes whether antecedents of anaphora are located within the same clause, in the previous clause, in the same paragraph or in another paragraph (Ariel 1990: 18f.). The *saliency* criterion evaluates whether a referent assumes the function of a main character in the discourse or not (Ariel 1990: 24f.). The parameter *unity* specifies the relationship between the clauses containing anaphora and their antecedents, i.e., whether they are contained in a single narrative frame or not (Ariel 1990: 26f.). Finally, the factor *number of competing referents* indicates the number of discourse entities that can potentially be designated by a given referential form (Ariel 1990: 28). On the basis of these four parameters, Ariel (1990) gradually distinguishes between referents of low and high accessibility. The shorter the distance to its antecedent, the more salient a referent, the closer the structural connection between anaphora and antecedent and the smaller the number of competing referents, the more accessible a discourse referent is (Ariel 1990: 18-30).

In the following, I will consider two of Ariel's (1990) parameters as potentially determining the referent chosen by *weeŋ*, namely the *distance* and *saliency* criteria.[15] Note that the distance between the anaphora and its antecedent is counted in words, in contrast to Ariel (1990) (who chose the clause), to account for more fine-grained differences. In addition, I will consider a third parameter, i.e., whether or not the referent takes on the role of subject in the preceding clause. As Diessel (1999) shows, demonstrative pronouns in German, which, from a grammatical point of view, can potentially denote several referents of a preceding sentence, denote a discourse unit that did not function as the previous subject but, for example, as a verbal object (Diessel 1999: 96).

In the following, three examples from the corpus are described and discussed, all of which contain either a demonstrative pronoun or a noun phrase containing *weeŋ* that can potentially denote multiple referents. After first describing the context and the competing referents of the three examples, in a second step the accessibility parameters will be applied to the competing referents. This will finally allow us to suggest which parameter(s) determine which referent is chosen by a form that is marked by *weeŋ*. The first example is shown in (10).

[15] The number of competing referents is irrelevant at this point, since this criterion cannot serve to compare the accessibility of referents potentially denoted by the same referential expression. The criterion of unity, on the other hand, is not applied, since a study on narrative boundaries is still pending. Currently, Hellwig & Schneider-Blum (in preparation) are working on this topic.

(10) *ʊ-kʊt-ɪ* *ɲ = ihinʌ,* *pɨnʌ* *u-pul-i-y = ii*
P-take-TR ERG = PRON3PL PRON3SG P-blow-TR-EP = APP

kɪ-ŋɛ *ii = c-ibʌ = yaa* *mak* *pɨnʌ* *ʊ-dʊ-ʊl,*
SG-mouth APP = SG-child = DEM2 then PRON3SG P-stop-MID

u-hweel = ii *mak* *pɨnʌ* *ʊ-dʊ-ʊl,*
P-whistle = APP then PRON3SG P-stop-MID

kɨcimbʌri = weeŋ *ɘ-daa-w-aa = ṭaŋ = ɪɪ* *tuŋkwiyʌʌk*
young.child = DEM3 P-run-EP-INS = LOC3P = APP SG.hat

ɪɪ = waṭɪŋ
APP = SG.owner
'They took [it] and one (of them) (child_2) whistled for that child (child_1), and then he (child_1) stopped. [He] (child_2) whistled for [him] (child_1) and he (child_1) stopped. That young child (child_2) ran with the hat to its owner (child_1).' (20190108_HamidPearFilm 022-024)

In (10), two children act as agents of the string of action. While one child, labeled 'child_1', lost his hat without noticing, the other child, labeled 'child_2', found it, together with other children. To return the hat, child_2 whistles for child_1, who has already left. Child_1 hears the whistle and stops. In the last sentence of this example, child_2 acts as the agent of the action again, as the return of the hat to its owner is described. In this clause, the noun phrase *kɨcimbʌriweeŋ* 'that child' could in principle refer to both children. However, *= weeŋ* signals that child_2 is referred to. The accessibility criteria of both potential referents of *kɨcimbʌriweeŋ* 'that child' can be seen in TABLE 2 and will be discussed later.

(11) *pɨnʌ* *dʊ-dʊwa* *yaya = yaŋ*
PRON3SG FUT-descend go.repeatedly = LOC3P

ɨ = k-ʌhunen = weeŋ; ***ku-weeŋ*** *ɘ-dɛɛk-aa-y = ɪɪ*
DIR = SG-woman = DEM3 SG-DEM3 P-scoop-INS-EP = APP

mɨnʌ *yaya = yaŋ*
PRON.ERG3SG go.repeatedly = LOC3P
'He later went repeatedly to that woman. That one (lit.: that one he) scooped water for [her] and went repeatedly to her.' (11_Hamad_4 154/155)

In (11), two protagonists are involved: the main protagonist of the narration, a boy, and a less salient female protagonist. In the first clause, the boy is referred to with a personal pronoun and described as repeatedly visiting the woman, who in turn is referred to with the external noun phrase *ɨkʌhunenweeŋ* 'to that woman'. The second clause begins with the demonstrative pronoun *kuweeŋ* 'that one' and refers to the boy, although either protagonist, i.e., the boy or the woman, could be referred to. Again, the accessibility criteria applied to both protagonists that could possibly be denoted by *kuweeŋ* are shown in TABLE 2.

(12) *ʊ-kʊna = a = ṭaŋ* *a = y-ɪhwaa;* *ɪhwaa = dɛ!*
P-ban = SOUR = LOC3P SOUR = EP-people [INTERJECTION]

pɨnʌ *maṭɛ* *cɪ* *rih-il-iŋ = dɛ,* *aaah*
PRON3SG and.then go turn-MID-VENT = REP aaah

ɪhwaa = weeŋ *ihinʌ* *an-da-y-ɪŋ,* *an-kʌmuh*
people = DEM3 PRON3PL 3PRF-run-EP-VENT 3PRF-give.up

pɨnʌ; *pɨnʌ* *tɪɪn* *twak*
PRON3SG PRON3SG ONLY [IDEOPHONE]
'[He] protected [them] from the people (i.e., the Wale); hey, people [addressing his audience]! he returned; ah, those people (Tima) they had run away, [they] left him behind, he is all alone.' (09_Hamad_2 032/033)

In example (12), the main protagonist, mentioned by name several clauses before, protects his kin from another ethnic group, the Wale people, as described in the first clause. Hereafter, this protagonist is referred to by the personal pronoun *pɨnʌ* 'he', while the following term *ɪhwaa* 'people' could essentially refer to either the Wale or the Tima people. In order to clarify that it is the protagonist's kin and not the recently mentioned Wale people who are being referred to, *= weeŋ* is attached to the noun. The accessibility criteria of both groups of referents can be seen in TABLE 2.

In TABLE 2, two parameters of accessibility (*distance* and *saliency*), as identified by Ariel (1990), as well as a third variable, i.e., whether a given referent fulfilled the role of the subject in the preceding clause, are applied to the competing referents potentially referred to by the noun phrases or pronouns in bold in examples (10) to (12). The referents to which the speakers actually refer are highlighted in grey. Which parameter (or parameters) seems to determine the identity of a form marked by *weeŋ* is discussed below.

EXAMPLE	COMPETING REFERENT	DISTANCE (IN WORDS)	SALIENCY	SUBJECT OF THE PRECEDING CLAUSE?
(10)	child_1	2	main character	yes
	child_2	5	side character	no
(11)	boy	4	main character	yes
	woman	1	side character	no
(12)	protagonist's people	16	side characters	no
	hostile people	6	side characters	no

TABLE 2: Accessibility criteria applied to the protagonists in (10)-(12)

First, the saliency of the referent is not considered to influence which of the competing referents is chosen. In both (10) and (11), one of the two referents is a main character in the narrative, while the other has to be categorized as a minor character. However, while, in (11), the main character, i.e., the boy, is referred to by the demonstrative pronoun *kuweeŋ* 'that one', in (10) the secondary character, i.e., child_2, is referred to by the noun phrase *kɨcimbʌriweeŋ* 'that young child'. In example (12), both competing protagonists are side characters, and neither of them was the subject of the preceding clause. Thus, the saliency of referents does not seem to influence which referent is denoted.

Second, the choice of referent does not depend on whether the referent functions as the subject of the preceding clause or not. This is shown again in examples (10) and (11). In both examples, one of the two referents potentially denoting the noun phrase or pronoun containing *weeŋ* fulfilled the role of the subject in the immediately preceding clause. However, whereas in (10) the referent that did not previously function as the subject is selected by the noun phrase *kɨcimbʌriweeŋ* 'that young child', in (11) the referent indicated by the demonstrative pronoun *kuweeŋ* 'that one' is the subject of the preceding clause. In example (12), none of the competing protagonists has functioned as subject before. Thus, the realization of *weeŋ* does not necessarily indicate a change of subject.

Finally, distance is the only parameter that can be used in all three examples to predict which referent is indicated. That is, noun phrases and pronouns containing *weeŋ* in all examples take up the referent that is textually more distant (and hence less accessible and prominent) than its competing referents, as can be seen in TABLE 2.

The function of *weeŋ* of resolving ambiguous references is attested 16 times (30%) in this corpus.

To conclude this section, it can be noted that the function of *weeŋ* presented above provides evidence for one of the potentially universal uses of demonstratives identified by Himmelmann (1996): the *tracking use* (Himmelmann 1996: 226). Demonstratives performing this function help the listener to "keep track of what is happening to whom" (Himmelmann 1996: 226). Specifically, noun phrases or pronouns that contain *weeŋ* were shown to refer to the textually more distant of several potential referents.

2.2 Marking referents known by personal knowledge

Another function of *weeŋ* can be described as the reference to entities that both interlocutors know outside the frame of the current discourse. As a result, these forms marked by *weeŋ* do not refer to textual antecedents. Three examples will be presented to illustrate this function. Before presenting an example from the narrative corpus, two elicited examples will be described. Evidence from elicited examples seems necessary because corpus examples that refer to extra-textually known referents cannot be identified unambiguously for two reasons. First, there is no certain proof that the interlocutors actually knew the referents marked with *weeŋ* outside the frame of the discourse. Although in most cases this can be easily determined from the narrative context (see the explanation of corpus example (15) below), the interpretation of the interlocutors' extratextual information about the referents remains somewhat speculative. Second, the corpus examples that seem to provide evidence for the marking of referents known by personal knowledge, and were thus exclusively assigned to the function presented here, often denote referents that have already been introduced in the text. Thus, in these cases, it cannot be ruled out with certainty that *weeŋ* actually serves to signal coreference to these textual antecedents, a function described in SECTION 2.1. We turn now to the evidence from elicitation.

In (13), a mother-tongue speaker of Tima was asked how the meaning of the noun phrase *ikihina ʊkwalɔŋ* 'this mountain (place)' changed if the proximate demonstrative clitic *=na* was replaced by *=weeŋ*. The speaker then produced the clause shown in (13), explaining that the addition of *=weeŋ* implies that the referent, i.e., the 'mountain', is known by both interlocutors, which in turn would not be implied if one of the other two demonstrative clitics were added.[16]

[16] HKD_20230129_metalinguistic-comment_01

(13) *[…]* ***ɨ=k-ihi=weeŋ*** *ʊ=k-walɔŋ=lɛɛy=ɪ*
DIR=SG-place=DEM3 DIR=SG-mountain=POSS1PL=SEL

ɪ-hɪɪ=nɛɛy=ɪ
PL-know=1PL.INCL=SEL
'[…] to the place of our mountain which we (incl.) know.'
(HKD_20230129_elicitation_01)

Similarly, in (14), the speaker refers to a man with a noun phrase marked with *=weeŋ* who has not been mentioned in the previous discourse. The demonstrative clitic is apparently added to express that the referent is known to both interlocutors.

(14) *wɔrt̪ɔmaadɔh=weeŋ* *nʌŋ* *t̪awʊ* *dʌŋ=ʌ* *k-ʌhu*
SG.man=DEM3 here Thawu like.this=FOC SG-name
'The man (who was) here is called Thawu.' (no recording)

Having shown that *weeŋ* serves to mark referents known through personal knowledge in the above elicitations, I will conclude by describing an example from the corpus to show how this function of the demonstrative is manifested in narratives. In (15), the entity *iihiyaa ɔkarkaman* 'to the places of the Karkaman' is referred to three times. First, the noun phrase *iihiyaa ɔkarkaman* 'the places of the Karkaman' introduces the location into the discourse.[17] Second, the location is referred to by *yaanʊŋ* 'there'. Third, the locative referent is taken up by the demonstrative pronoun *kuweeŋ* 'that (one)', which in turn introduces the relative-like clause *ɪyɔɔwaa pampaŋ mɪhɪ* '(where) we danced the drum dance in former times'.[18] The function of *kuweeŋ* 'that (one)' can be described as follows: it leads the addressee to identify the designated referent (the places of the Karkaman) on the basis of shared knowledge of an ancient tradition that took place at this location, i.e., a former dance place.

(15) *ɪ-cɪ* *ɨ=i-hi=yaa* *ɔ=karkaman* *yaanʊŋ*
P-go DIR=PL-place=DEM2 DIR=Karkaman there

ku-weeŋ *ɪ-yɔɔ-w-aa* *pampaŋ* *mɪhɪ*
SG=DEM3 1PL-dance-EP-INS SG.drum ancient
'He went to the places of the Karkaman there, that one (where) we danced the drum dance in former times.' (11_Hamad_4 142)

---

[17] The distal demonstrative clitic *=yaa* is seemingly attached to highlight the physical distance of the place referred to in relation to the speaker. It thus has exophoric reference.
[18] See Schneider-Blum's contribution to this volume for a discussion of whether Tima has relative clauses.

The function of *weeŋ* to refer to entities known outside the frame of the discourse is only attested three times (6%) in this corpus. This is probably related to the non-interactional nature of the narrations under study. One may hypothesize that *weeŋ* more frequently indicates personal knowledge of a referent when being used in spontaneous conversations. This remains to be investigated.

This use of *weeŋ*, i.e., the identification of a referent on the basis of personal knowledge, resembles the *recognitional use* of demonstratives, one of the demonstrative functions identified by Himmelmann (1996: 230-240). Himmelmann states that demonstratives in several languages fulfil this function as they draw on "knowledge that is assumed to be shared by the communicating parties due to a common interactional history or to supposedly shared experiences" (Himmelmann 1996: 233) in order to enable the identification of the referent by the addressee.

3 Discussion

In the following sections, I discuss the findings of the present study. Specifically, I will first give an overview of how the contexts of use of *weeŋ* resemble Himmelmann's (1996) potentially universal uses of demonstratives, followed by a discussion of what semantic-pragmatic feature the three uses of *weeŋ* have in common.

3.1 Consistencies and deviations from Himmelmann's (1996) taxonomy

This paper has examined the functions of the demonstrative *weeŋ*, identifying three uses: its function to signal the coreference of a noun phrase with a textually distant antecedent (SECTION 2.1.1), its contribution to the resolution of ambiguous references by marking the textually more distant of several competing referents (SECTION 2.1.2) and its function of marking that the referent denoted is known to interlocutors outside the frame of discourse (SECTION 2.2). These three uses of *weeŋ* provide evidence for two functions identified by Himmelmann (1996) as potentially universal uses of demonstratives, namely the *tracking use* and the *recognitional use*.

TABLE 3 shows which use of *weeŋ* corresponds to which of Himmelmann's potentially universal uses; specifically, the marking of entities known by personal knowledge corresponds to the recognitional use, while the marking of long anaphoric distance and the disambiguation of reference both correspond to the tracking use. TABLE 3 also shows how frequently these uses were attested in the present corpus. Note that only 47 of the 54 forms of *weeŋ* examined in the corpus are included in TABLE 3. This is because the remaining seven

instances performed functions not covered by those presented. What function they perform remains to be investigated.

FUNCTION OF *weeŋ*	NUMBER OF TOKENS	CORRESPONDING FUNCTION DESCRIBED BY HIMMELMANN (1996)
Marking of entities known by personal knowledge	3	recognitional use
Anaphoric long distance marking	28	tracking use
Marking the less activated of several possible referents	16	

TABLE 3: Functions of *weeŋ*

3.2 Common pragmatic profile

Regardless of whether *weeŋ* denotes referents that have textual antecedents or signals reference to entities known outside the frame of discourse, a common pragmatic profile can be observed, i.e., the marked referents are known to the addressee, but their identification is marked as difficult or, in other words, as requiring mental effort. Specifically, *weeŋ* always marks referents that either have low activation status in the listener's memory or have low activation in relation to other referents.

This proposed common pragmatic profile is supported by a common formal property: noun phrases or pronouns containing *weeŋ* are often realized in addition to other coreferential nouns. This tendency relates to the proposed common pragmatic profile. Since *weeŋ* qualifies referents as difficult to identify, further descriptions are provided to facilitate their identification. This link has already been described by Himmelman (1996), who observes the tendency of demonstratives performing the recognitional use to "incorporate anchoring or descriptive information [...] to make the intended referent more accessible" (Himmelman 1996: 230).

To be specific, one third of the elements marked with *weeŋ* function as appositions or dislocated topics; noun phrases or pronouns containing *weeŋ* are

used as appositions 20% of the time and as dislocated topics 13% of the time.[19] An example of a noun phrase containing *= weeŋ* and annotated as an apposition is shown in (16).

(16) *i-ṭibi-y-aa = ṭaŋ* *yala* *ʊ-wʊt-ɪ* *mɨnʌ*
P-fill-EP-INS = LOC3P come.on P-take-TR PRON.ERG3SG

ŋ = kʊlala = weeŋ *y-amʊh = ɛ = yɛ*
ERG = comrade = DEM3 PL-flour = FOC.PL = REP
'She filled it and then she, that comrade, took (it) as if (it were) flour.' (031007_Daldum_ClanDividing 015)

In (16), two sisters act as agents of the action. While one sister pours ashes into a pot, the other sister takes them. Since the reference of the personal pronoun *mɨnʌ* 'she' in the second clause is ambiguous, i.e., it could refer to either protagonist, the narrator adds the appositional noun phrase *ŋkʊlalaweeŋ* 'that comrade', which clarifies the identity of the chosen referent by pointing to the textually more distant antecedent.

Similarly, demonstrative pronouns whose root is *weeŋ* are attested as introducing relative-like clauses that help the addressee to identify a referent. Specifically, three of the four demonstrative pronouns containing *weeŋ* occur as arguments within a relative-like clause. An example is given in (17), where the demonstrative pronoun *iweeŋ* is an apposition of the preceding noun phrase *ibʌrimbʌriweeŋ* and serves as the subject of the relative-like clause *ʊkʊnɛṭaŋɪɪ*.

(17) *ibʌrimbʌri = weeŋ* *i-diik,* ***i-weeŋ***
young.children = DEM3 P-walk.away PL-DEM3

ʊ-kʊnɛ = ṭaŋ = ɪɪ
P-ban:TR = LOC3P = APP
'Those young children went (away), those (who had) helped him collecting.' (20190108_HamidPearFilm 020)

In (17), the narrator describes the action of a group of children who have already been introduced in the narrative. In the sequence immediately preceding (17), however, several other children were the protagonists of the action. The speaker

[19] For the purposes of this study, dislocated topics were identified as such if they fit one of the following two descriptions: either they are nominal elements realized within an intonation unit before the one containing their predicate, or they are coreferential with another nominal element closer to the predicate. Appositions, on the other hand, have been marked as such when they are realized in an intonation unit after the one containing the predicate on which they depend, or when they are realized after a nominal element with which they are coreferential.

thus uses the noun phrase *ibʌrimbʌriweeŋ* 'those young children' to refer to the textually more distant group of children. However, the narrator apparently found it necessary to additionally realize a relative-like clause in which he specifies which children are being referred to. The demonstrative pronoun *iweeŋ* 'those (who)' introduces this relative-like clause, herewith specifying the semantically specified group of referents.

4 Conclusion

This paper investigated the functions of the demonstrative *weeŋ* in Tima. Based on a qualitative corpus analysis of staged narratives, it was shown that three functions can be distinguished. First, the demonstrative marks referents that have textually distant referents (SECTION 2.1.1). Second, the demonstrative is used to resolve ambiguous references, where it picks up the textually most distant out of several competing referents (SECTION 2.1.2). Third, speakers use *weeŋ* to signal that a referent is known to the interlocutors outside the frame of the discourse (SECTION 2.2). Finally, it was shown that these functions have in common that referents marked with *weeŋ* are known to the addressee, but their identification is marked as difficult or, in other words, as requiring mental effort (SECTION 3.2).

Abbreviations

1	first person	INCL	inclusive
2	second person	INS	instrumental
3	third person	LOC	locative
AP	antipassive	LOG	logophoric
APP	applicative	LOW.TR	low transitivity
CAUS	causative	MOD	modifier
COND	conditional	OPT	optative
COP	copula	P	person
DEM1	first demonstrative (*nʌ/na*)	REP	reported
DEM2	second demonstrative (*yaa*)	PL	plural
		PLUR	pluractional
DEM3	third demonstrative (*weeŋ*)	POSS	possessive
		POT	potential
DIR	directional	PRF	perfect
EP	epenthetic element	PRON	pronoun
ERG	ergative	SEL	selective
EXCL	exclusive	SG	singular
FOC	focus	SOUR	source
FUT	future	TR	transitive
IMPFV	imperfective	VENT	ventive

References

Alamin, Suzan. 2012. *The Nominal and Verbal Morphology of Tima*. Cologne: Rüdiger Köppe.

Ariel, Mira. 1990. *Accessing Noun-Phrase Antecedents*. London: Routledge.

Bashir, Abeer. 2010. *Phonetic and Phonological Study of the Tima Language*. Khartoum: University of Khartoum dissertation.

Bühler, Karl. 1934. *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Fischer.

Becker, Laura & Gertrud Schneider-Blum. 2020. Morphological marking of contrast in Tima. *Glossa: A Journal of General Linguistics* 5(1). 1-35. https://doi.org/10.5334/gjgl.1098

Chafe, Wallace L. 1980. *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex.

Cornish, Francis. 2018. Anadeixis and the signalling of discourse structure. *Quaderns de Filologia: Estudis Lingüístics* 23. 33-57.

Diessel, Holger. 1999. *Demonstratives: Form, Function and Grammaticalization*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.42

Dimmendaal, Gerrit J. 2009. Tima. In Gerrit J. Dimmendaal (ed.), *Coding Participant Marking: Construction Types in Twelve African Languages*, 338-355. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.110

Dimmendaal, Gerrit J. 2014. The grammar of knowledge in Tima. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *The Grammar of Knowledge: A Cross-Linguistic Typology*, 245-259. Oxford: Oxford Academic. https://doi.org/10.1093/acprof:oso/9780198701316.003.0012

Dimmendaal, Gerrit J. 2018. Reconstructing Katloid and deconstructing Kordofanian. In Gertrud Schneider-Blum, Birgit Hellwig, & Gerrit J. Dimmendaal (eds.), *Nuba Mountain Language Studies – New Insights*, 383-416. Cologne: Rüdiger Köppe.

Dimmendaal, Gerrit J. & Gertrud Schneider-Blum. In preparation. *A Grammar of Tima*. Leiden/Boston: Brill.

ELAN (Version 6.4). 2022. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. https://archive.mpi.nl/tla/elan

Finkbeiner, Rita. 2018. Deixis und Anapher. In Frank Liedtke & Astrid Tuchen (eds.), *Handbuch Pragmatik*, 186-197. Stuttgart: Metzler. https://doi.org/10.1007/978-3-476-04624-6_18

Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274-307. https://doi.org/10.2307/416535

Haig, Geoffrey & Stefan Schnell. 2014. *Annotations Using GRAID (Grammatical Relations and Animacy in Discourse): Introduction and Guidelines for Annotators*. Version 7.0. https://multicast.aspra.uni-bamberg.de/

Hellwig, Birgit & Gertrud Schneider-Blum. In preparation. Split ergativity and larger discourse units in Tima (Niger-Congo, Sudan).

Heusinger, Klaus von. 2002. Specificity and definiteness in sentence and discourse structure. *Journal of Semantics* 2002 (3). 245-274. https://doi.org/10.1093/jos/19.3.245

Himmelmann, Nikolaus P. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.), *Studies in Anaphora*, 205-244. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.33.08him

Meerpohl, Meike. 2012. *The Tima of the Nuba Mountains (Sudan) – A Social Anthropological Study*. Cologne: Rüdiger Köppe.

Næss, Åshild, Anna Margetts & Yvonne Treis (eds.). 2020. *Demonstratives in Discourse (Topics at the Grammar-Discourse Interface 6)*. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.4054814

Peeters, David, Emiel Krahmer & Alfons Maes. 2021. A conceptual framework for the study of demonstrative reference. *Psychonomic Bulletin & Review* 28. 409-433. https://doi.org/10.3758/s13423-020-01822-8

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

RStudio Team. 2020. *RStudio: Integrated Development for R.* RStudio, PBC, Boston, MA. http://www.rstudio.com/

Schiborr, Nils N., Stefan Schnell & Hanna Thiele. 2018. *RefIND — Referent indexing in natural language discourse*. Unpublished.

Schneider-Blum, Gertrud. 2013a. *A Tima-English Dictionary*. Cologne: Rüdiger Köppe. Online version: http://tima-dictionary.mine.nu/

Schneider-Blum, Gertrud. 2013b. Personal pronouns in Tima. In Thilo C. Schadeberg & Roger M. Blench (eds.), *Nuba Mountain Language Studies*, 281-296. Cologne: Rüdiger Köppe.

Appendix

VARIABLE	DESCRIPTION	VARIANTS
dt_app	Does the noun phrase, including the demonstrative root, function as an apposition or as a dislocated topic?	no dislocated topic apposition
sel	Is the noun phrase, including the demonstrative root, marked by the selective clitic?	no selective marker
rel	Is the noun phrase, including the demonstrative root, further specified by a relative clause or itself part of a relative clause?	no followed by a relative-like clause inside a relative-like clause
ds	Is the noun phrase, including the demonstrative root, part of direct or indirect speech?	no direct speech indirect speech
endo_exo	Does the noun phrase, including the demonstrative root, refer to an entity on an exophoric or endophoric level?	endophoric exophoric unclear
distance	How far apart are the demonstrative marked anaphoric element and its antecedent?	number of clauses
ant_new	Was the antecedent of the demonstrative marked anaphoric element newly introduced?	no new antecedent old antecedent

TABLE 4: Variables and variants of the annotation