UD Annotation of Experience Clauses in Tigrinya

Nazareth Amlesom Kifle
Østfold University College
nazareth.a.kifle@hiof.no

Michael Gasser
Indiana University
gasser@iu.edu

Abstract

We are developing a treebank for Tigrinya within the Universal Dependency (UD) framework. UD proposes a set of universal grammatical relations such as `nsubj`, `obj` and `iobj` to capture dependency relations between words in any language. However, for some classes of verbs it is not a straightforward matter to know what grammatical relations the verbs are categorized for. In this paper we discuss the decisions we have had to make for the annotation of arguments of experience verbs in the Semitic language Tigrinya, which exhibit a number of unusual morphosyntactic properties. We describe a classification of experience verb roots in the language, based on the various ways in which the core experiencer and stimulus arguments are realized syntactically and morphologically and on which valence-changing operations the roots permit. We supplement our analysis with data from a morphologically analyzed Tigrinya corpus.

1 Introduction

We are developing a morphologically rich Universal Dependency (de Marneffe et al., 2021) treebank for the Semitic language Tigrinya. In addition to the extensions required to accommodate dependencies within as well as between words, we face several annotation challenges because of the mismatch between morphology and syntax and the unusual behavior of some verbs.

In this paper we focus on the category of experience verbs. Such verbs possess arguments that undergo some sort of mental, emotional or sensory experience and which exhibit variation in their morphosyntactic encoding in a wide variety of languages (Belletti and Rizzi, 1988; Næss, 2007; Pesetsky, 2000; Croft, 1993, 577-580). Experiencer predicates are typically categorized for an experiencer, the argument that experiences the mental state, and a stimulus, the argument that instigates the experience. Some predicates express the experiencer as a subject, while others express it as an object. This syntactic variation can be illustrated by the English predicates *fear* and *like*, on the one hand, and *frighten* and *please*, the other hand, where the experiencer corresponds to the subject 'I' in 'I fear snakes.', but with the object 'me' in 'Snakes frighten me'.

In this study we aim to outline a classification of experience verb roots on the basis of the syntax and morphology of their base forms as well as their passive and causative forms, where applicable. To our knowledge, there are no studies dedicated to Tigrinya experiencer verbs, only one that briefly describes the constructions (Kifle, 2011, 128-133). There is some work on experience verbs in the closely related language, Amharic (Amberber, 2005; Workneh, 2019) and, where relevant, we look at how Tigrinya clearly differs from Amharic.

In our study we rely not only on the native-speaker intuitions of one of us, but also on a morphologically analyzed corpus, which reveals statistical tendencies for particular roots and subcategories within the categories we propose.

This paper is divided into eight sections. Following this introduction, in section 2, we give a brief description of the morphologically enriched treebank we are developing. Section 3 presents a brief introduction to the morphosyntactic properties of Tigrinya, indicating in general how we annotate syntactic and morphological dependencies in our treebank. In Section 4, experiencer clauses are briefly described. Section 5 presents our morphologically analyzed corpus and the corpus data. Section 6 covers the method we used to categorize Tigrinya experience roots. In Section 7, we present the four categories of Tigrinya experience roots. Finally, in Section 8, we summarize our conclusions and outline future work to cover other possible arguments of experience verbs in the language and to automatically classify experience roots.

Proceedings of the Eighth International Conference on Dependency Linguistics (Depling, SyntaxFest 2025), pages 120–129

2 A morphologically enriched treebank

In the Tigrinya treebank we are creating, we segment morphologically complex words, treating all inflectional morphemes as tokens with their own parts-of-speech, lemmas, features, and dependencies. We do not separate derivational morphemes. We also maintain the distinction between subword tokens and morphologically complex words, making use of the CoNNL-U extension for handling multi-token expressions (https://universaldependencies.org/format.html#words-tokens-and-empty-nodes) for this purpose, as is done in the existing Amharic (https://universaldependencies.org/am/index.html) and Yupik (https://universaldependencies.org/ess/index.html) UD treebanks. Other UD treebanks that treat inflectional morphemes as tokens with their own relations to stems but do not maintain a separate word level that groups subword units together are the Beja (https://universaldependencies.org/bej/index.html) and Japanese (https://universaldependencies.org/ja/index.html) treebanks.

One of our goals in making relations explicit at both the morphological (within-word) and syntactic (between-word) levels is to explore and elucidate the complex ways in which participants are encoded both within a verb and as explicit nominals. The mapping between categories of pronominal affixes on verbs and case marking on nominals is not one-to-one in the language (Kifle, 2011, 66ff.). For example, the object pronominal suffix that typically marks definite accusative objects can also code applicative objects of intransitive verbs that are understood as affected participants which negatively experience the action of the verb. Moreover, pronominal suffixes serve as embedded pronouns instead of merely being agreement makers. In addition, the object case marker codes both accusative, dative and applicative objects. A further reason for segmenting verbs, nouns, and adjectives stems from our interest in using the treebank to train linguistically enriched language models and machine translation systems. The treebank will provide linguistically motivated subword units as an alternative to the segments generated by statistical methods such as byte-pair encoding that are the norm for such models (Gezmu, 2023).

3 Tigrinya morphosyntax

Tigrinya belongs to the family of Semitic languages spoken in Ethiopia and Eritrea. Like the other languages in this family, it is written in the Ge'ez abugida writing system. In our treebanks, we make use of Ge'ez orthography only, including for the morphological segmentation of words, but for the purposes of this paper, we add phonetic transcriptions and indicate the segmentation of words only when this is necessary to make a point.

The Ethiopian-Eritrean Semitic languages share many of the properties of other Semitic languages (e.g., template-based morphology, obligatory subject agreement, object agreement) as well as a number of properties of their own (e.g., verb final clause structure) (Feleke, 2021; Demeke, 2003; Hetzron, 1972). In this section, we describe morphological and syntactic properties of verbs and nominals that are relevant for the annotation of experience verbs and their arguments.

3.1 Verbs and valence-changing derivation

Tigrinya verbs consist of a stem and affixes coding subject and object agreement. Subordinate verbs take additional prefixes representing conjunctions and, for relative verbs, optional adpositions representing the case of the modified nominal.

(1) ስለዝረኣየቶ
sɨlə-zɨ-rəʔay-ət-to
since-that-see.PFV-SB3SF-OB1,3SM

'since she saw him'

As in other Semitic languages, verb stems are in turn derived from a root consisting of a series of consonants and a template consisting of a pattern of vowels inserted between the consonants and sometimes the gemination of one of the consonants. The language distinguishes four basic tense-aspect-mood categories, differing in their templates and their subject agreement affixes.

In addition to its base (simplex) form, each root can also appear in one or more derived forms, traditionally called *ʔaʕɨmad* (ኣዕማድ), corresponding to the *binyanim* of Hebrew and the *'awzaan* of Arabic verbs. Each *ʔaʕɨmad* has separate templates for each of the language's four tense-aspect-mood categories. As is usual for Arabic, we will refer to the different *ʔaʕɨmad* possibilities as "forms." For a given root, there may be as many as eight forms, in addition to the base form. In this paper we consider

only three of these: the BASE, the PASSIVE and the CAUSATIVE. Note that the specific interpretation of what we are calling PASSIVE and CAUSATIVE varies with the root. For example, some roots have no BASE form, and it is the PASSIVE or CAUSATIVE form that functions as the base form for these roots (Kifle, 2011, 61). We will refer to verb roots and stems using the 3rd person singular masculine perfective, as is conventional for Semitic languages.

3.2 Subject and object agreement

As in other Afro-Asiatic languages, verbs in Tigrinya are obligatorily inflected for subject person-number-gender agreement. In our morphologically enriched treebanks, we segment off subject agreement affixes and annotate the dependency joining the verb stem to them with the relation `nsubj`, adding the sub-relation `:aff` to distinguish them from the syntactic relations with the same label, as is done by Kahane et al. (2021, 51) for their morpheme-based treebank for Beja.

Tigrinya does not have a neuter gender, and 3rd person singular masculine (3SM) agreement is used to refer both to inanimate nouns that are lexically masculine and to unspecified dummy entities. As in modern Hebrew (Halevy, 2023, 10-12), it does not also have a locative or a demonstrative expletive, such as *there* in 'There is water in the glass.', or a dummy subject pronoun, such as *it* in 'It is hot', as in example (2).[1]

(2) ሞይቁ
moyq-u
be.hot.PFV-SB3SM

'It got/is hot.'

Verbs in Tigrinya may also take object agreement suffixes, also called "object suffix pronouns." These appear in two types, which we refer to as "object1" (OBJ1) and "object2" (OBJ2), following Kifle (2011, 104). The suffixes may refer to both objects (direct and indirect) and to applicative arguments, for example, *-to* (O1,3SM) in ርእያቶ *riʔiyat-to*, 'she saw him'; *-llu* (O2,3SM) in ርእያሉ *riʔiya-llu*, 'she saw for/on him'. While OBJ1 most often represents the direct or indirect object of a transitive verb, it may also represent a malefactive argument of an intransitive verb, for example, *-to* (O1,3SM) in ሞይታቶ *moytat-to*, 'she died on him/to his detriment.' (Kifle, 2007, 2011, p.119).

Each verb may take at most one object suffix; thus, speakers must choose between the objective and applicative suffixes when both are applicable to the arguments of a verb. In our treebank, we segment off the object suffixes. Since neither category of suffix corresponds directly to the UD `obj` relation, we annotate dependencies from the stem to the two types of suffixes with the special morphological relations `obj1:aff` and `obj2:aff`.

3.3 Nominals and case

Subjects in Tigrinya are not marked for case. Direct and indirect objects may take the objective prefix ን- *n-*.[2] Definite objects are obligatorily marked for case. The objective case marker also functions as the dative case marker (Kifle, 2007; Kievit and Kievit, 2009), marking arguments we annotate as `iobj`.

(3) a. አስቴር ንዮሴፍ ርእያቶ [Accusative]
ʔaster ni-yosef riʔiy-at-to
Aster OBJ-Yosef see.PFV-SB3SF-O1,3SM

'Aster saw Yosef.'

b. አስቴር ንዮሴፍ ህያብ [Dative]
ʔaster ni-yosef hiyab
Aster OBJ-Yosef gift
ሂባቶ
hib-at-to
gave.PFV-SB3SF-O1,3SM

'Aster gave Yosef a gift.'

Figure 1 shows the syntactic and morphological dependencies within sentence (3a). Co-referential nouns and verb affixes are indicated with the same color. Note that the co-reference relations are not explicit in the dependency graph.

The language has a set of adpositions that mark different semantic roles such as instrumental (ብ *bi*), locative (ኣብ *ʔab*), associative (ምስ *mis*) and elative (ካብ *kab*). As we shall see in some of the examples below, there is no simple one-to-one or one-to-many mapping between the categories of object affixes on verbs and the case markers and adpositions.

[1] We use the following abbreviations in interlinear glossing. 1: 1st person, 2: 2nd person, 3: 3rd person, AUX: Auxiliary, CAUS: Causative, DEF: Definite, F: Feminine, IPFV: Imperfective, M: Masculine, O1: Object1, O2: Object2, OBJ: Objective, PASS: Passive, PST: Past, PFV: Perfective, P: Plural, POSS: Possessive, PRS: Present, REL: Relative, S: Singular, SB: Subject.

[2] This prefix is normally referred to as "accusative," but we prefer "objective" because it can mark indirect as well as direct objects.

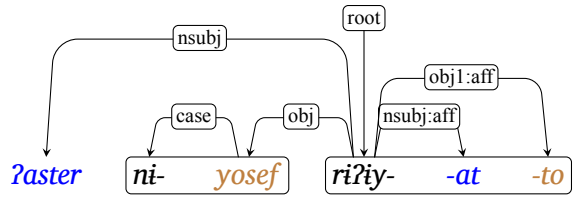

Figure 1: Dependencies in (3a). Blue tokens represent the subject, brown tokens the direct object.

4 Experience clauses

Experience clauses contain experiencer predicates, also known as "psychological predicates" (Postal, 1971, chapter 6) and "mental verbs (Croft, 1993, 55), that denote events that affect the consciousness of the experiencer such as its emotional or mental state or bodily sensation (Verhoeven, 2014, 130). Experience clauses are characterized by the presence of at least one of the two core semantic roles, the animate participant undergoing the experience, the EXP(ERIENCER), and the event or entity causing the experience, the STIM(ULUS) (Dowty, 1991; Croft, 1993; Klein and Kutscher, 2015). Languages have different means for leaving either the EXP or the STIM unspecified, for example, in the English sentences *this film is depressing*, which foregrounds the STIM, and *I'm depressed*, which foregrounds the EXP. As we will also see, experience predicates expressed by adjectives, such as *be quiet*, in English are normally expressed by verbs in Tigrinya.

With particular experience predicates, additional semantic roles are possible. Sometimes an external CAUSER needs to be distinguished from the STIM argument, for example, *news* in the sentence *the news made her dislike her teacher*.

Because experiencers can be perceived with different degrees of control over the experienced states and events, it is common in the world's languages for experience clauses to deviate from prototypical transitivity (Næss, 2007, 196). Experiencer nominals commonly appear as both subjects and objects, and when they are objects, they may be characterized by unusual case marking patterns.

5 Corpus Data

As we have access to a morphological analyzer for Tigrinya (https://github.com/hltdi/HornMorpho), we are able to assess how much information a morphologically analyzed corpus of sentences can provide about the statistical tendencies characterizing particular roots in the different categories of experience verbs we will be proposing.

First, we consider what morphological agreement features we expect for the EXP and STIM arguments. Experiencers are normally people, so all three persons, including in particular 1st and 2nd, should be possible features of the affixes agreeing with the EXP argument. Thus the absence of 1st and 2nd person agreement features for a particular affix can indicate that it does not refer to an experiencer. Stimulus features, on the other hand, are relatively unconstrained: experiences can be caused by people as well as inanimate objects and events. Impersonal verbs with "dummy" subjects are a special case; they always take 3SM subject agreement.

We can estimate a root's transitivity by looking at the proportion of instances that have an OB1 suffix, but this is only an indication of transitivity because (1) the suffix is only obligatory for definite objects and (2) though this is by far the most common use of the suffix, it can also function as a malefactive applicative agreement marker on intransitive verbs. Another measure of transitivity is the occurrence and frequency of the PASSIVE form of the root.

We ran a dedicated morphological analyzer on 1,000,000 Tigrinya sentences from the TLMD corpus (https://zenodo.org/records/5139094). For each verb root occurring in at least 10 unambiguous words, for each of the three forms under consideration, BASE, PASSIVE, and CAUSATIVE, we counted the occurrences of different subject and object agreement features.

For comparison we ran a morphological analyzer for the related Amharic language on 100,000 sentences from the CACO corpus (https://github.com/andmek/CACO).

6 Method

We start with the basic distinction between verbs taking EXP subjects and those taking EXP objects (Fleischhauer, 2016, 263-285). Because we are concerned with the annotation of the arguments of experience verbs, it is experience clauses, rather than simply experience verbs, that we will be discussing.

For each experience verb root that we consider, we will examine each of the three main forms that occur for that root: BASE, PASSIVE, and CAUSATIVE. For each of these forms, we will look

at how EXP and STIM are coded both morphologically and syntactically, and we will classify the roots on the basis of these properties. The result will be up to three morphosyntactic schemas for each root. For each schema we will be concerned with how the canonical roles are realized syntactically and morphologically and which UD relations we use for annotating each argument, both morphological agreement affixes and explicit nominal arguments of the verb.

7 Tigrinya Experience Verbs

The analysis of Tigrinya experiencer verbs reveals four categories which are outlined below.

7.1 Subject-experiencer verbs

Subject-experiencer (SE) verbs fall into two categories, intransitive verbs, which leave the STIM unexpressed, and transitive verbs, which code the STIM as direct object.

7.1.1 Intransitive SE verbs

Experiencer verbs such as ሰንበደ *sənbədə* 'be shocked', ሓዘነ *ħazənə* 'be sad,' and ዓገበ *ʕagəbe* 'be satisfied' are typical examples of the intransitive subject experiencer (ISE) class. The BASE form of this class is illustrated in (4).[3]

(4) አስቴር ሰንቢዳ [Tir]
ʔaster sənbid-a
Aster be.shocked.PFV-SB3SF

'Aster is shocked.'

ISE roots such as ሰንገደ *sənbədə* typically have a CAUSATIVE form in addition to the intransitive BASE SE form. The CAUSATIVE form takes the STIM as subject and the EXP as direct object. This is illustrated in (5).

(5) እቲ ወረ ንኣስቴር
ʔiti wərə ni-ʔaster
the news OBJ-Aster
ኣሰንቢዱዋ
ʔasənbid-u-wa
be.shocked.CAUS-SB3SM-O1,3SF

'The news shocked Aster.'

The fully segmented dependency tree for the sentence is shown in Figure 2.

[3] In all of our examples with an explicit EXP, this will be the feminine participant Aster.

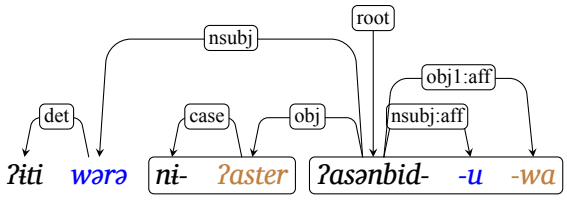

Figure 2: Dependencies in (5). Blue tokens represent the subject, brown tokens the direct object. Segmented words are surrounded by rectangles with rounded corners.

These CAUSATIVE forms can also appear without an explicit object EXP, where the focus is on the experience, independent of any particular EXP. We will refer to such clauses as "stimulus only" clauses. (6) is an example with the CAUSATIVE of the root *sənbədə* 'be shocked'.

(6) እቲ መብረቕ ኣጋዩ
ʔiti məbrəx' ʔaʒyu
the lightning very
የሰንብድ ነይሩ
yəsənbid nəyru
be.shocked.CAUS.IPFV.SB3SM AUX.PST

'The lightning was very shocking.'

The CAUSATIVE forms of these roots with relative subordinating morphology correspond to causative experiential adjectives in languages such as English: ዘሰንብድ *zəsənbid* be.shocked.REL.CAUS.IPFV.SB3SM 'shocking' (lit., 'that which causes shock').

ISE verb roots normally have no PASSIVE form. The exceptions are roots that lack a BASE form. For these verbs, the PASSIVE form behaves like the BASE form of a verb like ሰንበደ *sənbədə*, as in (4). Examples are the intransitive roots ተሓጐሰ *təħag^{w}əsə* 'be happy', ተጣዕሰ *təṭ'aʕlə* 'regret', and ተቖጥዐ *təx^{w}əṭ'ʕə* 'be angry.'

Corpus data for four ISE verbs confirm what we expected: that there is a frequent CAUSATIVE but no PASSIVE form, that OBJ1 suffixes are rare with the BASE forms but common with the CAUSATIVE forms, and that 1st and 2nd person subjects are frequent with the BASE form.

7.1.2 Transitive SE verbs

Tigrinya also has a set of transitive SE verbs (TSE) taking the STIM as direct object in the BASE form. Examples are ፈርሐ *fərħə* 'fear', ናፈቐ *nafəx'ə* 'miss', ሓፈረ *ħafərə* 'be embarrassed (over),' and

ጸልኣ *s'əliʔə* ‘hate.’ As expected for transitive verbs, these roots usually have PASSIVE as well as BASE forms. We annotate the EXP subjects of these PASSIVE verbs as `nsubj:pass`.

These roots also have derived CAUSATIVE forms related to the BASE forms in the manner of pairs like English *fear* and *frighten*. The CAUSATIVE of ፈርሐ *fərħə* ‘fear’ is illustrated in (7).

(7) ዮሴፍ ንኣስቴር
yosef *ni-ʔaster*
Yosef OBJ-Aster
ኣፍሪሑዋ
ʔafriħ-u-wa
fear.CAUS.PFV-SB3SM-O1,3SF

‘Yosef frightened Aster.’

The CAUSATIVE forms appear frequently in the stimulus only pattern, like the CAUSATIVE of intransitive SE verbs, as illustrated in (8).

(8) ዝብኢ የፍርሕ
zibʔi *yəfirriħ*
hyena fear.CAUS.IPFV.SB3SM
እዩ
ʔiyyu
AUX.PRS.SB3SM

‘A hyena is scary.’

The corpus data reveal that the roots in this class differ significantly with respect to transitivity, with ናፈቐ *nafəx'ə* ‘miss’ taking an OBJ1 suffix in 57% of the sentences in the BASE form, while this is true for only 21% of the sentences with the BASE form of ፈርሐ *fərħə* ‘fear.’ On the other hand, ፈርሐ *fərħə* ‘fear’ has a common PASSIVE form, whereas there are no instances of the PASSIVE form of ናፈቐ *nafəx'ə* ‘miss’ in the data. [4]

7.2 Object-experiencer verbs

Because EXPs are not prototypical agents and may be construed with varying degrees of control, they often appear as objects of different sorts and in many languages, for example, Icelandic (Barðdal, 1999), Faeroese (Barnes, 1986), and Greek (Landau, 2009), have quirky properties not characteristic of canonical transitive sentences.[5]

7.2.1 OE verbs with ambient stimuli

For one set of object-experiencer (OE) roots in the BASE form, Tigrinya shows a mismatch in case and pronominal marking: the EXP is treated morphologically as the object of the verb but syntactically it shows a split transitivity combining subject and object properties which according to Malchukov (2005) arises from a functional tension to foreground the most prominent argument, i.e. the experiencer.

In (9) the EXP, Aster is optionally marked with the objective case. When it appears as a bare noun, superficially like a canonical subject because it lacks the objective prefix that is normally obligatory for a definite direct object, but agrees with the 3rd person singular feminine OBJ1 suffix on the verb.

(9) (ን)ኣስቴር ጸሚኡዋ
ni-ʔaster *s'əmi-u-wa*
Aster be.thirsty.PFV-SB3SM-O1,3SF

‘Aster is thirsty.’

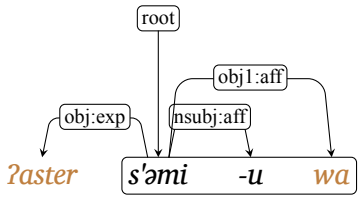

Figure 3: Dependencies in (9). Brown tokens represent the EXP.

Because the EXP nominals in such sentences are optionally marked with the objective affix and agree with the obligatory morphological object, we annotate them with the `obj` relation but add the sub-relation `:exp` to distinguish them from

[4] In fact the passive form of ናፈቐ *nafəx'ə* is possible in the language, for example, እቲ ዝሓለፈ ጊዜ ተናፊቑ አሎ። *ʔiti zihaləfə gize tənafix'u ʔallo* ‘The past time has been missed.’ This shows that we need to be cautious about concluding that a form is not possible simply because it fails to occur in the data.

[5] Landau (2009) identifies three types of languages based on a quirkiness scale. The first group comprises languages that have various options to code the EXP as dative, accusative and genitive quirky subjects, with Icelandic, Faeroese and Greek as typical examples. The second group allows only dative EXPs as subjects, with languages such as Italian, Spanish and Dutch showing this pattern. The third group does not allow quirky EXPs; that is, only nominative subjects can be used for EXP. Such languages include English, French and Hebrew. As we will see, Tigrinya is closest to the second group.

canonical direct objects, which require the objective prefix when definite. Roots of this type include ጸምአ *s'əmiʔə* 'be thirsty', ጠምየ *t'əməyə* 'be hungry', ደኸመ *dəxəmə* 'be tired', ጸምወ *s'əmməwə* 'feel lonely', ሰልቸወ *səlčəwə* 'be bored', and ሐመመ *ħaməmə* 'be sick, hurt'.

Morphologically, the subjects of these OE verbs are 3SM, similar to what Pesetsky (1995, 111) refers to as the unspecified stimuli behind "emotional weather," and, on the surface at least, identical to what Amberber (2005, 295), describing Amharic, calls "ambient causers". We will refer to clauses of this type as "ambient stimulus" (AS) clauses.

The picture is complicated by the fact that many of these roots also belong to the ISE category; for this reason Kifle (2011) treats them as applicative alternations. For example, the English gloss in (9) has another possible translation in Tigrinya, illustrated in (10).

(10) አስቴር ጸሚአ
ʔaster s'əmiʔ-a
Aster be.thirsty:PFV-SB3SF
'Aster is thirsty.'

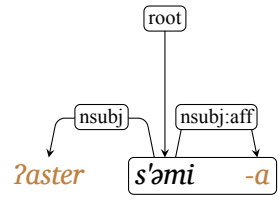

Figure 4: Dependencies in (10). Brown tokens represent the EXP.

In (10) the root ጸምአ *s'əmiʔə* behaves like an SE root; the verb's subject agrees with the EXP, Aster. In (9), on the other hand, the same root behaves like an ASOE root; the verb's object suffix agrees with the EXP and the EXP nominal has no case marker.

Examining the corpus data, we discover that the roots in this category differ strikingly with respect to their frequency of occurrence in the ISE and ASOE patterns. While the BASE forms ጸምወ *s'əmməwə* 'feel lonely' and ሰልቸወ *səlčəwə* 'be bored' have 3SM subjects with OBJ1 suffixes on the verb (indicating the ASOE pattern) in 40% and 82% of the instances, respectively, these proportions drop to 2.4% for ደኸመ *dəxəmə* 'be tired', 2.2% for ጠምየ *t'əməyə* 'be hungry', and 0.9% for ሐመመ *ħaməmə* 'be sick'.

Interestingly, two of the clearly related roots in Amharic exhibit quite different patterns: the proportion of 3SM subjects with object suffixes in the BASE form is 23% for ደከመ *dəkkəmə* 'be tired' and 90% for አመመ *ammməmə* 'be sick.' Another notable difference is that the Tigrinya roots in this category have no PASSIVE forms, while the PASSIVE forms for Amharic roots like አመመ *ammməmə* 'be sick' not only exist but are quite common.

For at least some of the roots that belong to both the ASOE and ISE categories, a further argument representing a generic stimulus is possible. In (11), ማይ *may* 'water' adds no information at all about the nature of the stimulus behind the state. Though not related etymologically to the verb, such an argument is analogous to "cognate objects" in other languages , for example *death* in *he died a peaceful death* (Austin, 1982; Jones, 1988; Pesetsky, 1995; Börjars and Vincent, 2008). We will refer to it as an "internal object."

(11) አስቴር ማይ ጸሚኡዋ
ʔaster may s'əmiʔ-u-wa
Aster water be.thirsty.PFV-SB3SM-O1,3SF
'Aster is thirsty (for water).'

At least for this root, the internal object is also possible when the root appears in the ISE pattern.

(12) አስቴር ማይ ጸሚአ
ʔaster may s'əmiʔ-a
Aster water be.thirsty.PFV-SB3SF
'Aster is thirsty (for water).'

We annotate the internal object as `obl:internal` in both (11) and (12).

Some of the roots in this category permit an explicit STIM argument that takes the form of the subject, so these then resemble the roots described in the next section. There is apparently a limited set of possible STIM subject arguments for these roots. With the BASE form of the root ደኸመ *dəxəmə* 'tire, be tired,' the noun ሓይል *ħayl* 'strength' with a possessive suffix is a common subject, as illustrated in (13). As in the AS pattern, we annotate the EXP in such sentences as `obj:exp` because it does not require the objective prefix when definite.

(13) አስቴር ሐይላ ደኽሚሙዋ
ʔaster ħayl-a dəxim-u-wa
Aster strength-her tire.PFV-SB3SM-O1,3SF

'Aster is tired.'

For the root ሐመመ *ħaməmə* 'sicken, be sick, hurt', the body part where the experience is centered may appear as the subject, as illustrated in (14). Again the EXP takes the form of an `obj:exp`, without the normal obligatory case marking.

(14) አስቴር ርእሳ ሐሙዋ
ʔaster riʔis-a ħim-u-wa
Aster head-her hurt.PFV-SB3SM-O1,3SF

'Aster's head hurts.'

For other roots in this category, if the speaker wants to refer to an explicit STIM, the CAUSATIVE form must be used, with the EXP in the form of a canonical object, that is, with an OBJ1 agreement suffix on the verb and the obligatory objective prefix on the nominal if definite. We annotate the EXP argument as `obj`. Sentences in this pattern may also include the internal object, which we annotate as `iobj:internal` in this case. See (15), in which ጨው *č'əw* 'salt' is the `nsubj`, አስቴር *ʔaster* is the `obj`, and ማይ *may* 'water' is the `iobj:internal`.

(15) እቲ ጨው ንኣስቴር ማይ
ʔiti č'əw ni-ʔaster may
the salt OBJ-Aster water
አጽሚኡዋ
ʔas'mi-u-wa
thirsty.CAUS.PFV-SB3SM-O1,3SF

'The salt made Aster thirsty (for water).'

7.2.2 OE verbs with explicit stimuli

For other OE roots, such as ገረመ *gərəmə* 'surprise, be surprising,' an explicit STIM subject is possible with the root's BASE form. This is illustrated in (16).

(16) ስርሑ ንኣስቴር
sirħu ni-ʔaster
action.his OBJ-Aster
ገረሙዋ
gərim-u-wa
surprise.PFV-SB3SM-O1,3SF

'His action surprised Aster.'

These roots may also appear in the ambient stimulus pattern, in which case the EXP, if definite, no longer requires the objective prefix.

(17) አስቴር ገረሙዋ
ʔaster gərim-u-wa
Aster surprise.PFV.SB3SM.O1,3SF

'Aster is surprised.'

With such roots, the BASE form may also appear in the stimulus only pattern, as illustrated in (18).

(18) ስርሑ ይገርም እዩ
sirħu yigərrim ʔiyyu
action.his surprise.IPFV.SB3SM AUX.SB3SM

'His action is surprising.'

Other roots in this category include ጨነቐ *č'ənəx'ə* 'worry,' ጸገመ *s'əggəme* 'trouble', and ሃወኸ *hawwəxə* 'disturb.'

Not surprisingly, the roots in this category have a PASSIVE form, in which the EXP is the subject and the STIM is an `obl`, as in (19) .

(19) አስቴር ብትዕግስቱ
ʔaster bi-tiʕgistu
Aster by-patience.his
ተገሪማ
təgərrim-a
surprise.PFV.PASS-SB3SM-O1,3SF

'Aster is surprised by/with his patience.'

But many of the roots also have a CAUSATIVE form, in which the EXP and STIM are realized as with the BASE form, as object and subject respectively. See 20, in which the subject, ትዕግስቱ *tiʕgistu*, functions as STIM.

(20) ትዕግስቱ ንኣስቴር
tiʕgistu ni-aster
his.patience OBJ-Aster
አገሪሙዋ
ʔagərrim-u-wa
surprise.PFV.CAUS-SB3SM-O1,3SF

'His patience surprised Aster.'

There are cases where the BASE and CAUSATIVE forms of such roots are interchangeable but others in which the experience CAUSER and STIM are separated, as in (21). In these cases the CAUSATIVE

form of the root is required, the CAUSER is the `nsubj`, and the STIM is realized as an `obl` argument.

(21) ዮሴፍ ንኣስቴር ብትዕግስቱ
yosef *ni-aster* *bi-tiʕgistu*
Yosef OBJ-Aster by-his.patience
ኣገረሙዋ
ʔagərrim-u-wa
surprise.PFV.CAUS-SB3SM-O1,3SF

'Yosef surprised Aster with his patience.'

7.3 Summary of categories

Here we summarize the morphosyntactic schemas we have described for experience clauses in Tigrinya. Syntactic and morphological relations are separated by a slash when they are different. To simplify, the morphological subrelation `:aff` is not included.

- Subject Experiencer: Intransitive
  Example: ሰንበደ *sənbədə* 'be alarmed'
  - Base: ሰንበደ *sənbədə*
    - EXP: `nsubj`
  - Causative: ኣሰንበደ *ʔasənbədə*
    - EXP: `obj/obj1`
    - STIM: `nsubj`
- Subject Experiencer: Transitive
  Example: ጸልኣ *sʼəlʔə* 'hate'
  - Base: ጸልኣ *sʼəlʔə*
    - EXP: `nsubj`
    - STIM: `obj/obj1`
  - Passive: ተጸልኣ *təsʼəlʔə*
    - STIM: `nsubj`
  - Causative: ኣጽልኣ *ʔasʼliʔə*
    - EXP: `obj/obj1`
    - STIM: `nsubj`
- Object Experiencer: Ambient Stimulus
  Example: ጸመወ *sʼəmməwə*'feel lonely'
  - Base: ጸመወ *sʼəmməwə*
    - EXP: `obj:exp/obj1`
  - Causative: ኣጸመወ *ʔasʼəmməwə*
    - EXP: `obj/obj1`
    - STIM: `nsubj`
- Object Experiencer: Explicit Stimulus
  Example: ገረመ *gərəmə*'surprise'
  - Base: ገረመ *gərəmə*
    - EXP: `obj/obj1`
    - STIM: `nsubj`
  - Passive: ተገረመ *təgərrəmə*
    - EXP: `nsubj`
  - Causative: ኣገረመ *agərrəmə*
    - EXP: `obj/obj1`
    - STIM: `nsubj`

8 Conclusions and Future Work

Our investigation has uncovered two categories within each of the Subject Experiencer and Object Experiencer classes of experience verbs in Tigrinya, each defined by a schema for each of the two or three forms that occur for the roots in the category. We have also seen that many roots belong to more than one category. In particular, roots such as ደኽመ *dəxəmə* 'tire, be tired' occur in both the Ambient Stimulus Object Experiencer and Intransitive Subject Experiencer categories. We might guess that the use of these roots in the Subject Experiencer pattern implies a more active role for the EXP, something we plan to explore in future work. We have also learned that specific roots may permit arguments that are not possible with others in the same category, for example, body part subjects with ሓመመ *ħaməmə* 'hurt' and internal objects with ጸምኣ *sʼəmʔə* 'be thirsty'.

We have not exhausted all of the possibilities for the arguments of experience predicates in the language. For example, the verb ሓመመ *ħaməmə* 'be sick, hurt' can take a malefactive argument representing a participant who is harmed by the experiencer's pain or ailment. In future work we plan to investigate the morphosyntax associated with these arguments and propose UD relations for them.

Finally, the existence of morphological data on thousands of Tigrinya roots opens up the possibility of classifying experience roots on the basis of their similarity to the categories we have outlined in the paper.

References

Mengistu Amberber. 2005. Differential subject marking in Amharic. In Mengistu Amberber and Helen de Hoop, editors, *Competition and Variation in Natural Languages: The Case for Case*. Elsevier.

Peter Austin. 1982. Transitivity and cognate objects in australian languages. In *Studies in transitivity*, pages 37–47. Brill.

Jóhanna Barðdal. 1999. The dual nature of Icelandic psych-verbs. *Working papers in Scandinavian syntax*, 64:79–101.

Michael Barnes. 1986. Subject, nominative and oblique case in Faroese. *Scripta Islandica*, 37:13–46.

Adriana Belletti and Luigi Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language & Linguistic Theory*, pages 291–352.

Kersti Börjars and Nigel Vincent. 2008. Objects and obj. In *Proceedings of the LFG08 Conference*, pages 150–168.

William Croft. 1993. Case marking and the semantics of mental verbs. In *Semantics and the Lexicon*, pages 55–72. Springer.

Marie-Catherine de Marneffe, Christopher D. Manning, Joakim Nivre, and Daniel Zeman. 2021. Universal Dependencies. *Computational Linguistics*.

Girma Awgichew Demeke. 2003. *The clausal syntax of Ethio-Semitic*. Ph.D. thesis, University of Tromsø.

David Dowty. 1991. Thematic proto-roles and argument selection. *language*, 67(3):547–619.

Tekabe Legesse Feleke. 2021. Ethiosemitic languages: Classifications and classification determinants. *Ampersand*, 8:100074.

Jens Fleischhauer. 2016. *Degree Gradation of Verbs*. Düsseldorf University Press.

Andargachew Mekonnen Gezmu. 2023. *Subword-based neural machine translation for low-resource fusion languages*. Ph.D. thesis, Otto-von-Guericke-Universität Magdeburg.

Rivka Halevy. 2023. Non-subject oriented existential, possessive and dative-experiencer constructions in modern hebrew–a cross-linguistic typological approach. *STUF-Language Typology and Universals*, 76(4):545–585.

Robert Hetzron. 1972. *Ethiopian Semitic: studies in classification*. 1. Manchester University Press.

Michael Allan Jones. 1988. Cognate objects and the case-filter. *Journal of Linguistics*, 24(1):89–110.

Sylvain Kahane, Martine Vanhove, Rayan Ziane, and Bruno Guillaume. 2021. A morph-based and word-based treebank for Beja. In *20th International Workshop on Treebanks and Linguistic Theories*.

Dirk Kievit and Saliem Kievit. 2009. Differential object marking in tigrinya. *Journal of African Languages and Linguistics*, 30(1):45–71.

Nazareth Amlesom Kifle. 2007. Differential object marking and topicality in Tigrinya. In *Proceedings of the LFG*, volume 7, pages 5–25.

Nazareth Amleson Kifle. 2011. *Tigrinya Applicatives in Lexical Functional Grammar*. Ph.D. thesis, University of Bergen.

Katarina Klein and Silvia Kutscher. 2015. *Lexical economy and case selection of psych-verbs in German*. Institut für Deutsche Sprache, Bibliothek.

Idan Landau. 2009. *The locative syntax of experiencers*. MIT press.

Andrej Malchukov. 2005. Split intransitives, experiencer objects and'transimpersonal'constructions:(re-) establishing the connection.

Åshild Næss. 2007. *Prototypical Transitivity*. John Benjamins.

David Michael Pesetsky. 1995. *Zero syntax: Experiencers and cascades*. 27. MIT press.

David Michael Pesetsky. 2000. *Phrasal movement and its kin*. MIT press.

Paul Martin Postal. 1971. *Cross-over phenomena*. Holt, Rinehart and Winston.

Elisabeth Verhoeven. 2014. Thematic prominence and animacy asymmetries. Evidence from a cross-linguistic production study. *Lingua*, 143:129–161.

Desalegn Workneh. 2019. *The Voice System of Amharic*. Ph.D. thesis, Arctic University of Norway.