Lang Resources & Evaluation (2020) 54:97–121
https://doi.org/10.1007/s10579-019-09465-5

ORIGINAL PAPER

Building the first comprehensive machine-readable Turkish sign language resource: methods, challenges and solutions

Gülşen Eryiğit[1] · Cihat Eryiğit[1] · Serpil Karabüklü[2,3] · Meltem Kelepir[2] · Aslı Özkul[2,4] · Tuğba Pamay[1] · Dilara Torunoğlu-Selamet[1] · Hatice Köse[1]

Published online: 17 April 2019

Abstract This article describes the procedures employed during the development of the first comprehensive machine-readable Turkish Sign Language (TiD) resource: a bilingual lexical database and a parallel corpus between Turkish and TiD. In addition to sign language specific annotations (such as non-manual markers, classifiers and buoys) following the recently introduced TiD knowledge representation (Eryiğit et al. 2016), the parallel corpus contains also annotations of dependency relations, which makes it the first parallel treebank between a sign language and an auditory-vocal language.

This research is supported under the project "A Signing Avatar System for Turkish to Turkish Sign Language Machine Translation" by The Scientific and Technological Research Council of Turkey (TUBITAK) with a 1003 Grant (No. 114E263) and under the project "Sign-Hub" by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 693349.

The convention in sign linguistics is to use the acronyms of sign languages as they are used by the Deaf community, namely, with the capital letters of the sign language name in the local spoken language. Thus, TiD represents the first letters of the Turkish words Türk İşaret Dili 'Turkish Sign Language'.

✉ Gülşen Eryiğit
gulsen.cebiroglu@itu.edu.tr

Cihat Eryiğit
ceryigit@itu.edu.tr

Serpil Karabüklü
skarabuk@purdue.edu

Meltem Kelepir
meltem.kelepir@boun.edu.tr

Aslı Özkul
asli.ozkul@bilgi.edu.tr

Tuğba Pamay
pamay@itu.edu.tr

Keywords Turkish sign language · TiD · Parallel dependency treebank · Turkish · Machine-readable · Parallel corpus

1 Introduction

Parallel texts are crucial resources for statistical machine translation (MT) studies, the vast majority of which have been conducted between auditory-vocal language pa irs having a writing system. The parallel data sets compiled from such languages are mostly obtained by the automatic alignment of written translated text materials. Even though written materials do not represent spoken language accurately, most MT systems are based on pairs of written language corpora since developing machine-readable speech data, which would heavily rely on speech recognition systems, is more costly and more prone to errors (Koehn et al. 2007). MT studies focusing on sign languages face an additional challenge: using written text for sign languages is troublesome since sign languages do not have conventionalized writing systems that would lead to large sets of written language materials. Thus, researchers developing an MT system where sign languages are involved must first take on the task of developing a data set which contains the written representations of the sign language data. One way of doing this is to annotate the video recordings of sign language data manually, the other is to employ a sign language recognition system that converts the visual form of the sign language to a written form. Manual annotation is inevitably a costly and slow procedure, and thus does not result in the creation of large data sets that can serve as the input for the training of statistical MT systems. Sign language recognition systems, on the other hand, do not yet yield high performances, again mostly due to the lack of training data. Therefore, despite their drawbacks, manually created parallel data sets are still very valuable resources for MT systems for the following reasons:

- As test data sets, they serve as important resources for intrinsic evaluation of MT systems using only written representations, leaving out errors resulting from speech or sign language recognition/generation components of an end-to-end system.
- Together with formal sign language representation definitions, they serve as prototypes for end-to-end systems, exemplifying how sign language recognizers

Dilara Torunoğlu-Selamet
torunoglud@itu.edu.tr

Hatice Köse
hatice.kose@itu.edu.tr

[1] Department of Computer Engineering, Istanbul Technical University, Istanbul 34469, Turkey

[2] Department of Linguistics, Boğaziçi University, Istanbul 34342, Turkey

[3] Present Address: Department of Linguistics, Purdue University, West Lafayette, USA

[4] Present Address: English Language Teacher Education, Istanbul Bilgi University, Istanbul, Turkey

should produce machine-readable outputs to feed MT modules or how MT modules should produce their outputs to successfully feed sign animation environments (e.g. avatars, humanoid robots).

Most of the sign language corpus studies mentioned in the literature have been conducted for the purpose of linguistic research [e.g. Leeson et al. (2006) for Irish Sign Language, Schembri et al. (2013) for British Sign Language, Johnston (2016) for Australian Sign Language, Wallin and Mesch (2015) for Swedish Sign Language, Koizumi et al. (2002) for Japanese Sign Language, Prillwitz et al. (2008) for German Sign Language, Crasborn and Zwitserlood (2008), De Vos et al. (2015) for Sign Language of the Netherlands and Özsoy et al. (2013) for Turkish Sign Language], and the annotation schemes developed for those corpora are not easily transferable to machine translation studies. The annotations in such corpora are mostly designed to take human-readability and understandability into consideration and, as a result, are most of the time not machine-readable. Johnston (2008) is the first study to introduce ID-glosses and one of the preliminary studies towards machine readability of sign language corpora. In the few corpus studies which focus on machine translation, the compiled parallel data sets are generally small and domain specific, such as aeronautics and weather news [e.g. Su and Wu (2009) between Chinese and Taiwanese Sign Languages, Bungeroth and Ney (2004) and Bungeroth et al. (2006) for weather reports in German and German Sign Language, Bungeroth et al. (2008)'s ATIS corpus for air travel domain available in English, German, Irish Sign Language, German Sign Language and South African Sign Language]. Othman et al. (2012) proposes a rule-based translation method from English to American Sign Language in order to produce a large parallel corpus to be used in statistical studies.

The advantages of using syntactic information in rule-based machine translation have been reported (Galley et al. 2004, 2006; DeNeefe et al. 2007). Hence, an important resource for rule-based MT is parallel treebanks, in which parallel data sets collected from the languages in focus are syntactically annotated (Cuřín et al. 2004; Cmejrek et al. 2005; Megyesi et al. 2008; Ahrenberg 2007; Uchimoto et al. 2004; Tinsley et al. 2009). Since the 90s, dependency analysis (Tesnière 1959) has become a widely used approach for natural language parsing and as a result, dependency treebanks have become crucial resources. The most important reason for this popularity is that dependency trees provide reasonable approximation to the semantic layer and are directly usable in higher level NLP tasks.[1]

Sign languages used in different countries/communities differ substantially from each other (and also from the spoken languages used in these countries) at the lexical, morphological and syntactic levels. Due to recent developments in natural language technologies, the need for computerized processing for TiD has increased. The first machine-readable knowledge representation for TiD introduced in Eryiğit et al. (2016) makes available the necessary infrastructure for its utilization in resource creation: an interactive online dictionary platform and an

[1] The Swedish Sign Language Corpus Project (2017) and Östling et al. (2017) present the first dependency treebank for a sign language (Swedish Sign Language).

ELAN[2] (Wittenburg et al. 2006; Crasborn and Sloetjes 2008) add-on for TiD. This article introduces a bilingual TiD lexicon and the first machine-readable parallel treebank for the TiD - Turkish language pair using the above-mentioned knowledge representation and infrastructure. This resource has advantages over other TiD corpora intended for use in computerized studies due to its comprehensive annotation scheme, which includes sign language specific features (such as non-manual markers, classifiers and buoys).[3] Our study not only introduces a comprehensively annotated machine-readable corpus for TiD but also the first parallel dependency treebank between a sign language and an auditory-vocal language.

The article is structured as follows: Sects. 2 and 3 provide the methodology of data collection and annotation respectively, Sect. 4 discusses the challenges faced and the solutions provided, Sect. 5 gives the evaluation of the developed resource, and discussions and Sect. 6 concludes the article.

2 Methodology of data collection

In this section, we describe the data sources of the parallel corpus (between Turkish and TiD) that are the foundations of the introduced language resource (TiDLaR). The linguistic data of TiDLaR is restricted to the vocabulary items and sentences in the first grade primary school coursebooks (Dalkılıç and Gölge 2013; Boz et al. 2013; Demiroğlu and Gökahmetoğlu 2013) published by the Ministry of Education in Turkey, since this study was part of a project which was the first attempt to develop an MT system as a supplementary tool for primary school education of deaf and hard of hearing students.[4]

TiDLaR consists of two components that interact with each other: a lexical database of 1530 signs (collected under 990 entries) and a parallel corpus of 420 annotated Turkish-TiD sentence pairs. The lexical database (henceforth, Lexicon with a capital L) has been designed in such a way that it can both function as an independent dictionary and also serve as the lexicon of the automated translation system. In this article, we describe some relevant properties of this database as it functions as the lexicon of the annotated language resource. For a detailed description of the properties of the used online dictionary platform, the reader is referred to Eryiğit et al. (2016) which gives a survey of the available Turkish-TiD dictionaries and the reasons why a more comprehensive dictionary is needed. From now on in this article, we refer to this component as the (TiDLaR) Lexicon.

[2] ELAN (EUDICO Linguistic Annotator) is a professional tool for the creation of complex annotations on video and audio resources and is widely used for sign language annotation. There exist also other sign language annotation platforms such as iLEX (Hanke and Storz 2008) and SignStream (Neidle et al. 2001).

[3] Camgöz et al. (2016) introduces a sign language recognition corpus consisting of TiD signs and phrases from health and finance domains and Selçuk-Şimşek and Çiçekli (2017) a parallel dataset solely depending on word order correspondences between TiD and Turkish.

[4] This MT system is from written Turkish to avatar animated TiD.

2.1 Turkish-TiD parallel corpus (TiDLaR)

In order to make the collected dataset more suitable for machine translation studies (especially for the one focused on our above-mentioned project), we opted for eliciting the TİD utterances as translations of written Turkish sentences instead of alternative data collection approaches such as via visual stimuli. The chosen approach is very similar to the data collection processes employed for machine translation between auditory-vocal languages (see details in Sect. 1). This approach is preferred due to the facts that (1) not every sentence can be elicited with visual stimuli and (2) elicitation of sentences with a method other than translation would easily lead to TiD sentences whose meanings are completely different from the meanings of the Turkish counterparts, which would make the dataset hard to use in machine translation.

A total of 306 Turkish sentences in TiDLaR are from the first grade Social Studies (Dalkılıç and Gölge 2013) and Mathematics (Boz et al. 2013) coursebooks. The TiD sentences are created as translations of these written Turkish sentences. In some cases more than one possible translation for a Turkish sentence was offered by the TiD consultant. In those cases, all the proposed alternatives in TiD were recorded and annotated. The corpus thus contains the videos and the annotated files of 420 TiD sentences. Translation was carried out by a number of fluent signers of TiD. Two of these signers were fluent, non-native signers who became deaf in early childhood and were exposed to Turkish Sign Language before the age of 12. Both are active members of the TiD signing community in Istanbul. The third fluent signer is a professional interpreter, trained and employed by the Turkey Ministry of Family and Social Policies. She is a hearing child of Deaf[5] adults (CODA)[6] and thus, learned TiD as her mother tongue. Due to budget restrictions only one signing language consultant could be employed at a time. Therefore, the Deaf signers contributed to this study in a sequential fashion: the second one checked the translations of the first one and modified them when necessary. The hearing interpreter, on the other hand, both checked the translations of the first and the second one and also discussed with the second one the cases in which the signed translations were not clear.

Two hearing linguistics researchers fluent in Turkish Sign Language worked with the Deaf consultants in order to translate the written Turkish sentences to TiD. First, the consultant was asked to read the book on her own, and then the linguistics researchers went over the sentences with her in TiD to clarify the meanings of the sentences. It is crucial to note that Turkish is a second language for these Deaf consultants, and they are not very fluent readers of Turkish. Thus, in a number of cases, the researchers had to explain the meanings of the sentences to the consultants in order to elicit the best translations. Other challenges of translation

[5] The convention in sign language and Deaf studies is that the adjective Deaf (with capital D) is used when it refers to the community, culture and signers who identify themselves as part of the Deaf community culturally. The adjective deaf (with small d) is used for the medical condition.

[6] CODA stands for Children of Deaf Adults. This acronym is used in sign language and Deaf studies to identify this special population. CODAs are special in that they usually are brought up as bilinguals: they can speak both the sign language of their parents and the local spoken language.

will be discussed in Sect. 4.1. When the discussion and clarification of the meanings of the Turkish sentences were completed, the Deaf consultants were asked to translate these sentences into TiD as naturally as possible. They were asked to make sure that the translations did not look like signed Turkish, i.e. word-to-sign translations. In fact, when the second Deaf consultant and the professional interpreter watched the videos of the translated TiD sentences, they were also asked to judge these in terms of naturalness. Once the translation phase was completed, the TiD data were annotated using ELAN. The details of the annotation procedures are explained in Sect. 3.

2.2 TiDLaR lexicon

The TiD lexical items within the Lexicon come in the following categories: (1) one or more counterparts of the Turkish vocabulary items or multi-word expressions found in the sentences from the Social Studies (Dalkılıç and Gölge 2013), Mathematics (Boz et al. 2013) and Turkish (Demiroğlu and Gökahmetoğlu 2013) coursebooks, and (2) those signs that were uttered during the translation process from Turkish to TiD but do not have counterparts in the books. For instance, if the consultant reported that there are two signs for the color 'green' in TiD, both signs were recorded and entered into the Lexicon. Furthermore, if the consultant reported that there are two different signs for different (but related) meanings of a Turkish lexical item, then both of these signs were again recorded and entered. One such example is *çalışmak* 'to work' in Turkish. This verb is the counterpart of both 'to study' and 'to work' and these two different meanings are expressed with different signs in TiD (as in English).

Since sign languages do not have commonly used writing systems, it is common practice to use "glosses" in the local spoken language to represent the meanings of signs (Miller 2001). Thus, ID-glosses (Johnston 2008, 2016) in Turkish, following universal conventions in linguistic studies, are used to uniquely identify the signs in the TiDLaR Lexicon and Corpus. For each lexical entry, one or more of the following are provided[7]: (1) a video in which the lexical item is signed by a Deaf consultant, (2) if the sign is a noun and can be inflected for plurality, the video of its plural form, and (3) if there is another sign with the same meaning, the video of that sign.[8]

3 Methodology of annotation

The annotation of TiDLaR required different levels of expertise. Similar to the data collection stage, several annotators with different backgrounds played a role during the annotations. Our two Deaf annotators, who worked at the translation stage, also worked to transcribe the TiD utterances (provided in the TiD Tier under ELAN tiers

[7] See Eryiğit et al. (2016) for a detailed description of the annotation scheme of TiD signs.

[8] Plural formation in sign languages does not always involve simple concatenative inflection, and the form of the plurals of signs may depend on a number of factors (Kubuş 2008; Steinbach 2012).

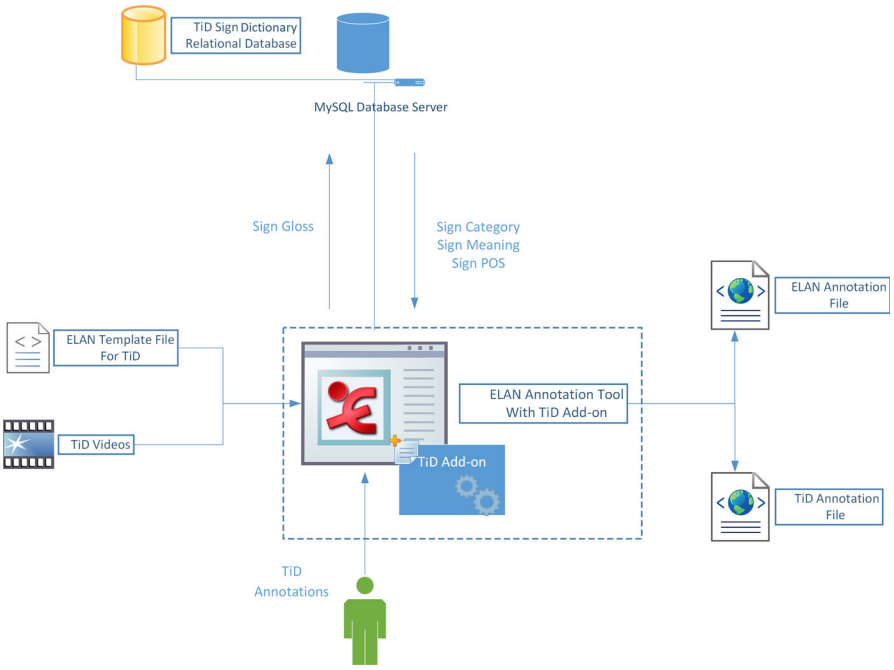

Fig. 1 TiD annotation architecture (Eryiğit et al. 2016)

Fig. 2). Then, two of our linguistics researchers fluent in TiD annotated the TiD utterances at the morphology layer. Since the dependency annotations also require a different level of expertise, two other annotators experienced in Turkish morphology and syntax annotations (but with only a basic understanding of TiD) worked for the dependency annotation of both the TiD utterances and the Turkish sentences.

The annotation (knowledge-representation) scheme that we followed for TiD was previously introduced in Eryiğit et al. (2016). Similarly, for the annotation of the Turkish sentences, we followed the scheme introduced in Sulubacak et al. (2016b). Thus, for a detailed description of features and categories that constitute these annotation schemes, we refer the reader to these previous studies. In this article, after providing brief summaries of these annotation schemes, we focus on how we created this resource, our annotation procedure, the challenges we encountered during the work, and the places where we had to extend the TiD scheme.

The sub-sections below describe the annotation procedure used in this study to annotate the TiD sentences using ELAN extended with a TiD add-on described in Eryiğit et al. (2016) (Fig. 1) and the annotation procedure used to annotate the Turkish sentences using the ITU Treebank Annotation Tool (Eryiğit 2007b). The TiD add-on produces an extra XML file for the machine-readable TiD transcriptions (to be used in machine translation studies) while leaving the original ELAN XML file intact. By doing so, all the annotations made on ELAN tiers are still readable by

classical ELAN versions whereas the TiD annotation XML files together with the TiD schema file may be parsed by XML parsers and used in machine translation.

3.1 Annotation of the TiD utterances

Although there exist environments such as ELAN for the manual annotation of sign language video resources, these have been generally used in linguistic research with the aim of creating human-readable annotations. Recently, Eryiğit et al. (2016) has proposed a machine-readable representation scheme for TiD which is linked to ELAN manual annotation layers via a new add-on that enables manually-annotated TiD corpora to serve as sources for machine translation research (henceforth TiD add-on). This add-on mainly provides the following advantages: (1) it automatically extracts the sign annotations available from the Lexicon; (2) it provides the tools for annotating additional context-dependent features for signs (e.g. annotating the agreement information of agreeing verbs), and (3) it enables the annotator to enter detailed information about the signs that are context-dependent and not present in the Lexicon (e.g. classifiers, buoys). This add-on integrates a new TiD tab (cf. the top right of Fig. 2) that makes the TiD add-on features available.

A typical annotated ELAN file in our study contains six tiers: "Discourse", "Utterance", "Turkish", "TiD", "MainFlow" and "SupportFlow".[9] "Discourse" and "Utterance" tiers contain information related to the current annotation (e.g. an identification number in our case). Although the Discourse tier may contain one or more utterances, due to the structure of our course books and also for simplicity of annotation, we had only one utterance annotated per discourse in TiDLaR. The "Turkish" tier contains the Turkish sentence taken from the course book. The "TiD" tier contains the glosses of the TiD signs in the TiD sentence. This tier was annotated by a Deaf annotator.[10] The "MainFlow" and "SupportFlow" tiers contain the ID-glosses of the signs (specified with SMALL CAPS henceforth), which are most of the time retrieved from the Lexicon. Exceptional cases to this are explained in later sections. The ID-glosses were entered into and retrieved from the Lexicon by the linguistics researchers.

The MainFlow is the sequential realization of sign instances and the SupportFlow is the flow of the extra signs made by the non-dominant hand (Eryiğit et al. 2016). In other words, the MainFlow tier contains the representation of the signs which are produced either by only the dominant hand (the preferred hand of a signer while articulating one-handed signs) or by both hands. For example, if a sign is two-handed as in HOUSE, it is annotated only in the MainFlow. The SupportFlow tier,

[9] ELAN TiD Tier hierarchy is built on "included in", "time subdivision" and "symbolic subdivision" stereotypes as exemplified in Fig. 2. The reader may refer to ELAN guidelines http://www.mpi.nl/corpus/manuals/manual-elan_ug.pdf for further details on tier stereotypes.

[10] Some Turkish sentences were difficult to translate into TiD. For instance, the sentence "How did you express this feeling of yours?" was not possible to translate directly to TiD since the Deaf consultants reported that there are no signs for the notions "feeling" and "express". It was translated as: "How was it? Tell me." In such cases, the TiD tier contains this later translation as well appended to its glossing within square brackets (e.g. Fig. 7).

on the other hand, mainly contains the representations of signs produced only by the non-dominant hand. These are usually buoys and classifiers (see Sect. 3.1.1).[11,12]

The content of the TiD tab depends on (i.e. changes dynamically according to) the selected ELAN tier. The annotation features that become editable in the TiD tab when either the Discourse or the Utterance tier is selected are those that affect the entire sentence. When the MainFlow or the SupportFlow tier is active, the annotations at sign-level become available under three TiD sub-tabs: "Basic", "Arguments" and "Advanced" as was shown in Fig. 2 (emphasized by red rectangles on the figure). The remainder of this section focuses on our annotations at sign-level, utterance-level, and syntactic-level.

3.1.1 Sign-level annotations

The annotation procedure of sign-level annotations starts with retrieving the ID-gloss of each sign from TiDLaR Lexicon (see Fig. 2 Gloss field under the Basic Tab). Next, the main category of the sign is selected (e.g. lexical, classifier, or buoy) and annotations start at different levels. According to the selected category information, the related arguments are dynamically enabled for human annotation (under the Arguments sub-tab). If a sign in an utterance that is being annotated displays properties different from or additional to the ones already specified for it in its Lexicon entry, these are also specified (under the Advanced sub-tab). These properties, labeled as "non-manual markers", "plural", "incorporation", "incorporation type", and "hand usage", will be discussed below.

Lexical signs:

If the sign is a lexical sign, then its available meanings and parts of speech (e.g. noun, verb, or pronoun) are retrieved again from the Lexicon and the relevant tags in the context of the given TiD utterance are selected by our annotators (see Fig. 2, Parts of speech and Meaning fields under the Basic Tab). As specified above, the related arguments are dynamically enabled for human annotation according to the selected POS and category information. For instance, if the sign is defined as an agreement verb, information on its inflectional features (i.e. what it agrees with) is provided at this stage.

Classifiers:

In addition to lexical signs, the utterances in the corpus also contain context-dependent signs such as classifiers and buoys. Classifiers are manual signs that do not have specific lexical meanings but represent entities by denoting their most

[11] In contrast to many other sign language annotation conventions (Crasborn et al. 2015; Johnston 2016), in our annotation scheme, manual signs are not annotated based on whether they are articulated with the left/right hand or with the dominant/non-dominant hand. Therefore, two-handed signs, for instance, are annotated only on the MainFlow. We adopted this approach in order not to lose the atomicity of a sign for machine-readability purpose.

[12] It should be noted that it was not uncommon in the data for buoys and classifers to be signed with the dominant hand. In those cases, these signs were annotated in the MainFlow.

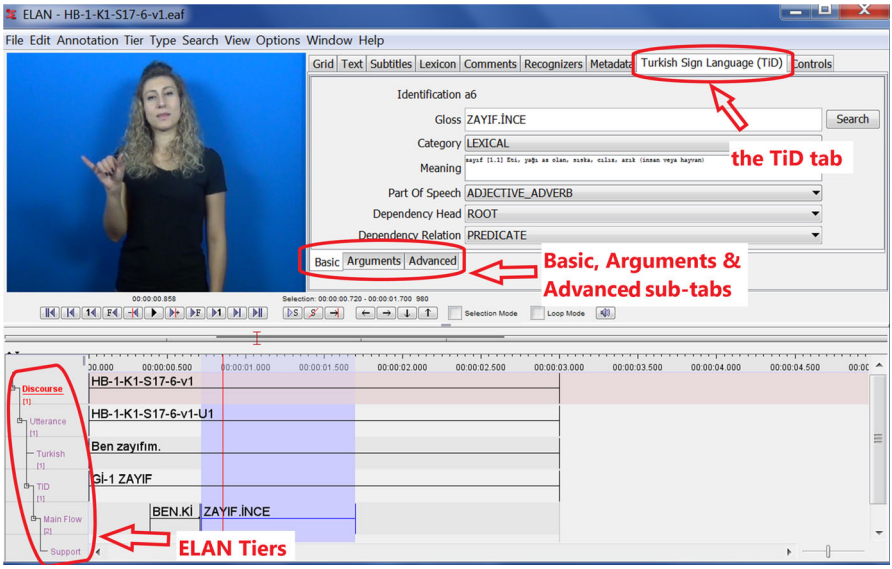

Fig. 2 Tiers and Tabs in ELAN TiD Add-on English glosses and translation: I THIN –> "I'm thin"

salient characteristics, such as a two-legged animal or a flat surface. They may represent entities in a static position as in "A cup stands on the table", or in motion as in "A person is walking".[13]

Since classifiers are mostly context-dependent signs (Zwitserlood 2012), it is not possible to enter all possible classifier variations in the Lexicon. That is why classifiers are annotated in the arguments tab by choosing their handshape, orientation, and movement, if any. To be able to track more than one classifier construction in an utterance, their glosses are given based on their semantic reference, such as PICTURE.CL, which has CL to distinguish it from the lexical PICTURE and signals its dependency relation (for later stage annotations, to be discussed below), differently than the lexical sign.

Following the TiD representation scheme, the classifiers are annotated under four classifier sub-categories: handling, body-part, entity, and size and shape. Handling classifiers represent entities that are being held and/or moved by an agent, such as representing the hand of an agent opening a door (Zwitserlood 2012). Body-part classifiers represent the body itself, such as eyes, hands or feet. Entity classifiers represent entities by denoting particular semantic and/or shape features such as using the index finger handshape to represent an upright human being. Size and

[13] Classifiers are iconic signs; however, iconic representation of an entity with a classifier may change from context to context, and from language to language (Perniss et al. 2010; Zwitserlood 2012). In other words, different classifiers may represent the same entity in different contexts. For instance, PENCIL could be expressed both with an entity and a handling classifier. If the handshape is index finger selected, the index finger represents the pencil as an entity in the context. On the other hand, if the handshape is baby-O, then it represents holding the pencil.

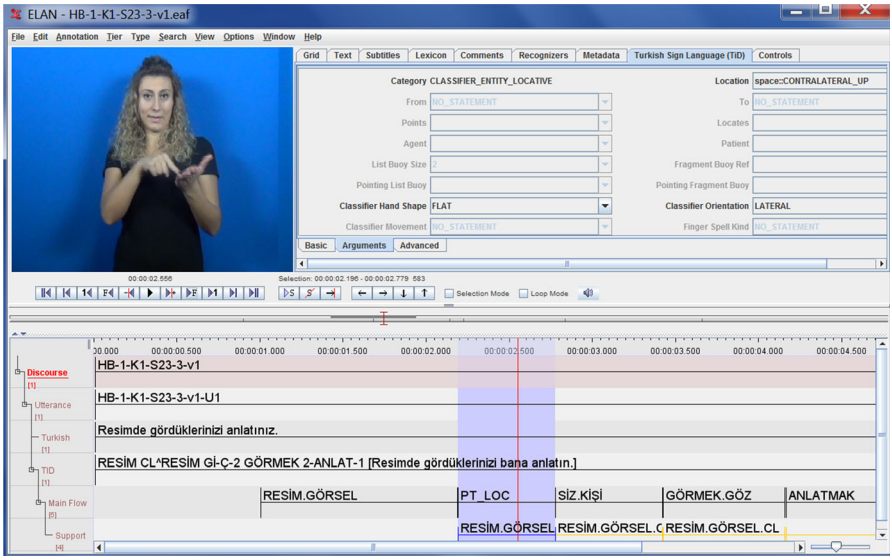

Fig. 3 Arguments tab when a classifier is annotated Eng.: PICTURE CL^PICTURE YOU SEE 2-EXPLAIN-1 –> "Explain to me what you see in the picture"

shape specifiers denote nouns according to the visual-geometric features of their referents. An example would be tracing the edges of a rectangular surface.

The classifier arguments (handshapes and movements) are taken from Kubuş (2008)'s exhaustive list. For instance, during the annotation of an entity-movement classifier, the annotator can choose among different movement patterns such as swimming, walking, sliding, going upstairs etc. In the case of an entity-location classifier, for instance, its handshape (e.g. flat), its locus in the signing space (e.g. contralateral-up) and its orientation (e.g. lateral) are specified, as shown in Fig. 3.

Buoys:

Buoys are signs that help guide the discourse by serving as conceptual landmarks as the discourse proceeds (Liddell 2003). That is, a buoy signals to the addressee that the conversation is related to the topic it represents. These signs are produced frequently by the non-dominant hand in a stationary configuration as the dominant hand continues producing the other signs. Some buoys appear only briefly, whereas others may be maintained during a significant stretch of signing. They can also be referred to (e.g. pointed at) by the dominant hand. Buoys are annotated under two types: list buoys and fragment buoys. A list buoy is used as enumeration when referring back to a list of items. The fingers of the non-dominant hand represent an ordered set of topics in the discourse. The signer may start with expressing the entire list, for instance, holding open four fingers and referring back to the first, second, third and the fourth finger in an order, or he/she may enumerate the list without first expressing the total number of the items in the list. A sign is called a fragment buoy when it associates the meaning of a sign with all or part of its final state of production. Fragment buoys are not lexicalized; they are created spontaneously.

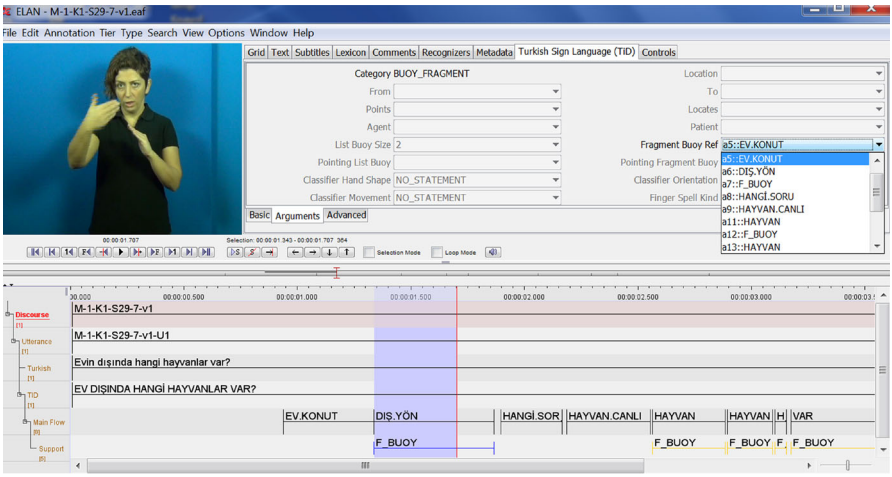

Fig. 4 The selection of Fragment Buoy reference for the sign EV.KONUT 'house' Eng.: HOUSE OUTSIDE WHICH ANIMAL THERE.ARE –> "Which animals are there outside the house?"

During the annotation in the Basic tab, one of these types is specified for a buoy sign.

The following example illustrates the annotation of fragment buoys. The first sign in the MainFlow in Fig. 4 is EV.KONUT 'house'  . After signing it, the signer keeps a part of the sign HOUSE on the non-dominant hand while signing DIŞ.YÖN 'outside' with the dominant hand. DIŞ.YÖN is annotated in the MainFlow and the buoy is annotated in the SupportFlow. When the category of a sign is selected as "fragment buoy", the annotator also indicates the fragment of which lexical sign functions as a buoy. In the example, the fragment buoy is linked to the lexical sign EV.KONUT 'house' via "Fragment Buoy Ref." field. A fragment buoy can also be referred back to by a pointing sign. In those cases, the category of the pointing sign is selected as POINTING_BUOY_FRAGMENT. Which fragment buoy it points to is also specified in "Pointing Fragment Buoy" field as an argument.

Non-manual markers:

Non-manual markers (NMMs) are gestures that have linguistic functions (e.g. facial expressions, eye brow position, head position, head nod or head shake, body lean, body tilt, mouth gestures and mouthings) (Pfau and Quer 2010). NMMs can be lexical features of individual lexical items and can thus occur only during the articulation of the manual sign. They can also have syntactic and pragmatic functions and thus can spread over larger constituents such as the verb phrase or the entire sentence. These functions can be marking topics, yes/no questions, negation, etc. In the TiDLaR annotation scheme, if a sign has a lexical non-manual marker, this is specified in its lexical entry when it is entered into the Lexicon. However, if that sign is signed with a different non-manual marker in the annotated utterance (e.g. with "topic"-marking non-manuals), then this additional non-manual marker is specified under the "Non-manual markers" feature in the Advanced tab.

Plurality:

Kubuş (2008) and Zwitserlood et al. (2012) report that TiD does not have a single systematic plural marker but employs different strategies such as reduplication and use of quantors. Thus, it is not possible to have a morphological rule that can predict the plural form of a given singular noun. The annotators choose from among the 26 plural formation strategies (such as "locative reduplication") reported in Kubuş (2008) when annotating the pluralization strategy of a plural noun.

Incorporation:

The feature "Incorporation" refers to cases where a sign co-occurs with another sign that is phonologically dependent on it.[14] The types of incorporation choices available here are "numerical incorporation", "negative incorporation" and "general incorporation". It is commonly observed in sign languages that number signs (1–5 or 1–9) incorporate into nouns, pronouns and temporal expressions (Costello et al. 2017). This is illustrated below. Figure 5a shows the number sign ONE in isolation. Figure 5b, on the other hand, shows a complex sign, meaning 'first grade', where ONE is incorporated into the sign GRADE. ONE in isolation is signed vertically in the neutral signing space with the index handshape. When incorporated, however, it is signed in the location of GRADE, which is the upper part of the arm of the non-dominant hand.

Hand usage:

The articulation of a sign may display different handedness properties than those specified for it in the Lexicon. Such deviations are annotated:

- if a sign is articulated with the non-dominant hand in the annotated utterance, whereas its entry in the Lexicon specifies its articulation with dominant hand;
- if a sign is articulated with both hands in the annotated utterance, whereas its entry in the Lexicon specifies its articulation with only the dominant hand;
- if a sign is articulated with a single hand in the annotated utterance, whereas its articulation is specified as two-handed in the Lexicon.

3.1.2 Utterance-level annotations

The properties annotated at the utterance-level are "Features", "Utterance Non-Manuals" and "Non Manuals Suppressed". "Features" refers to different clause types such as declarative, interrogative and exclamative, tense types such as past tense, aspect types such as perfective, and modality types such as obligation. "Utterance Non-manuals" refers to the non-manual markers that spread over the entire utterance, such as those marking a sentence as a yes/no question (see 3.1.3 for an explanation of NMMs). The information on the lexical non-manuals of the sign, if there is any, is available in the Lexicon and is retrieved together with other kinds of information of the sign. However, these lexical NMMs can sometimes be

[14] Note that this use of the term "incorporation" differs from its common use in theoretical linguistics where it is interpreted as a morpho-syntactic operation that combines at least two syntactic heads into a complex word.

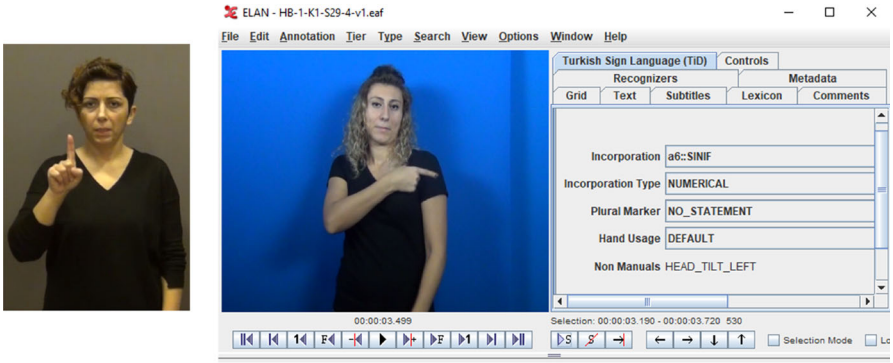

Fig. 5 Annotation of number incorporations **a** ONE **b**FIRST^GRADE

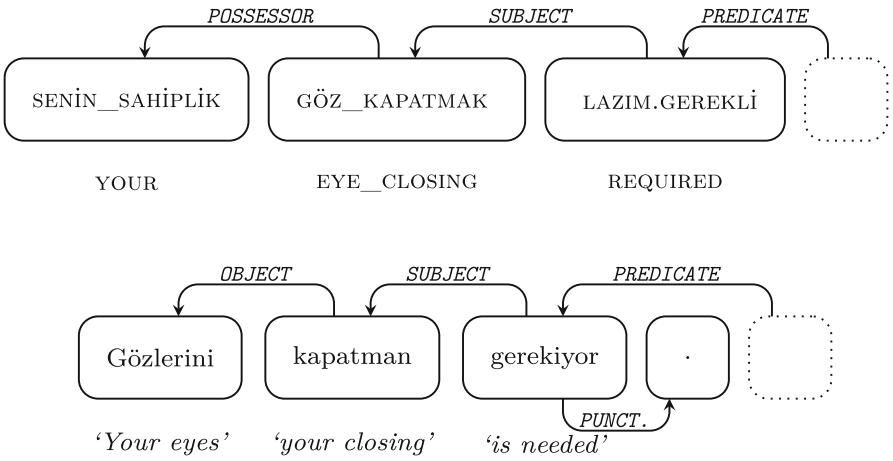

Fig. 6 Dependency graphs for a TiD utterance and its Turkish counterpart

suppressed by the utterance NMMs. In such cases, the suppression is annotated by selecting the suppressed non-manual under the "Non Manuals Suppressed" feature. The annotators annotate each utterance by first selecting the Utterance tier and then choosing from the relevant property values that become available on the TiD tab.

3.1.3 Dependency annotations

According to the dependency formalism, each syntactic token within a sentence should have been connected to one head token, and head tokens may have many dependent tokens linked to them. Following this rule, the head sign of each sign within the syntax tree of the parent utterance is annotated together with its dependency relationship type.

An example dependency tree is provided in Fig. 6 for a TiD utterance (meaning 'You should close your eyes.') and its Turkish counterpart (*Gözlerini kapatman gerekiyor.*). The annotator selects one of the sign glosses from a combobox as the dependency head of the current sign (e.g. GÖZ__KAPATMAK 'eye closing' is the dependency head of the sign SENİN__SAHİPLİK 'your') and a dependency relation (e.g. *POSSESSOR*) from a predetermined list of possible relation types (to be detailed in Sect. 4.2.2). A dummy ROOT node is added to each utterance in order to act as the root of the dependency tree depicted with an empty dashed rectangle in Fig. 6.

3.2 Annotation of the Turkish sentences

The studies for Turkish treebanking started in the early 2000s (Atalay et al. 2003; Oflazer et al. 2003) and have continued since then (Eryiğit 2007a; Sulubacak and Eryiğit 2013; Pamay et al. 2015; Eryiğit et al. 2015; Sulubacak et al. 2016a, b; Sulubacak and Eryiğit 2018). Following earlier studies, the Turkish part of the TiDLaR Corpus was annotated in two main layers: morphology and syntax. The new web version of the ITU Treebank annotation tool (Eryiğit 2007b; Pamay et al. 2015), used in recent studies, was selected as the annotation environment, and two experienced annotators worked during the annotation process. In the morphology layer, the annotators were supposed to choose a correct morphological analysis for each input Turkish token from an automatically produced list of possible analyses (Şahin et al. 2013; Eryiğit 2014). In the syntax layer, the dependency annotations were carried out according to Sulubacak et al. (2016b).[15] Each of the two annotators annotated half of the sentences and controlled the other half. In case of conflict, the annotators worked together and corrected the analysis upon their agreement.

Table 1 provides the main parts-of-speech (POS) tags and dependency relation distributions within Turkish and TiD sections of the TiDLaR corpus. The minor POS-tag and different non-manual type distributions in the TiD corpus are provided in the Appendix. As stated previously, during the annotations of the TiD section, the representation scheme provided in Eryiğit et al. (2016) was adopted. However, during the syntactic annotations, we needed new dependency relation types in addition to the basic ones given in this scheme. The challenges we faced and the newly added relation types are discussed in Sect. 4.2.

4 Challenges and solutions

Creation of the parallel corpus raised a number of challenges. The problems that emerged during the translation phase and the methods to ensure accuracy and reliability in translation that the researchers developed are described in the first sub-

[15] Although a very recent study (Sulubacak et al. 2016a) focuses on mapping the Turkish dependency grammar to Universal Dependencies (Nivre et al. 2016), we preferred to follow Sulubacak et al. (2016b) because of our annotators' experience on this framework, and we left mapping to the Universal Dependencies to future work.

Table 1 TiDLaR parts-of-speech tag and dependency relation distributions

POS Tag	Turkish	TiD	Dependency relation	Turkish	TiD
ADJECTIVE	176 (9.4%)	–	*ARGUMENT*	46 (2.5%)	45 (2.7%)
ADVERB	98 (5.2%)	–	*CONJUNCTION*	30 (1.6%)	27 (1.6%)
CONJUNCTION	41 (2.2%)	29 (1.7%)	*COORDINATION*	31 (1.7%)	74 (4.4%)
DETERMINER	10 (0.5%)	–	*DERIV*	176 (9.4%)	–
INTERJECTION	9 (0.5%)	–	*DETERMINER*	26 (1.4%)	30 (1.8%)
NOUN	719 (38.4%)	585 (34.9%)	*INTENSIFIER*	11 (0.6%)	2 (0.1%)
POSTPOSITION	47 (2.5%)	–	*MODIFIER*	442 (23.6%)	450 (26.8%)
PRONOUN	109 (5.8%)	273 (16.3%)	*MWE*	97 (5.2%)	47 (2.8%)
PUNCTUATION	304 (16.3%)	–	*OBJECT*	136 (7.3%)	134 (8.0%)
VERB	357 (19.1%)	316 (18.8%)	*POSSESSOR*	103 (5.5%)	148 (8.8%)
ADJ_ADVERB	–	201 (12.0%)	*PREDICATE*	300 (16.0%)	303 (18.1%)
ADPOSITION	–	72 (4.3%)	*PUNCTUATION*	304 (16.3%)	–
NUMERAL	–	25 (1.5%)	*RELATIVIZER*	1 (0.1%)	–
PARTICLE	–	16 (0.9%)	*SUBJECT*	148 (7.9%)	288 (17.2%)
POINTING	–	57 (3.4%)	*VOCATIVE*	19 (1.0%)	15 (0.9%)
BUOY	–	36 (2.1%)	*CASE*	–	17 (1.0%)
FINGER SPELL	–	6 (0.4%)	*NEGATIVE*	–	14 (0.8%)
CLASSIFIER	–	61 (3.7%)	*COREFERENCE*	–	39 (2.3%)
			CONTINUE	–	44 (2.6%)
Total	1870 (100%)	1677 (100%)	*TOTAL*	1870 (100%)	1677 (100%)

section below. The challenges in ELAN annotations and identification of dependency relations are more analytic in nature, and this issue is discussed in the following sub-section.

4.1 Challenges in translation: ensuring reliability

Needless to say, translating sentences of a spoken language into a sign language is very different from translating between two spoken languages that are used in educational settings, and the task thus comes with its own challenges. As explained in 2.2, TiD lexical items in the Lexicon of our Resource mostly consist of translations of words found in the first-grade course books. The task of translating these words into TiD was far from trivial. This is because TiD is not formally used in educational settings (Dikyuva et al. 2017) and the TiD lexicon lacks many signs that express the concepts found in school course books. This issue has been addressed in our study in the following ways: (1) if there was a chapter in the book with a large number of concepts whose counterparts are not in the TiD lexicon, the sentences in those chapters were not translated and left out of the sentential database, (2) the second Deaf signer consulted other signers in the community and the professional interpreter of the project to see whether there were signs used for

these concepts, however rarely, and (3) in some cases, the lexical items of Turkish were finger-spelled.

As mentioned in Sect. 3.1, once the initial translation of the sentences into TiD and the recording of the citation forms of the individual lexical items was completed by the first Deaf consultant, the second Deaf consultant watched each TiD sentence and the lexical item video and evaluated them for naturalness and accuracy. In a number of cases, she suggested a modification, and the new form was recorded and the initial version replaced by the new form. In another set of cases, when the second consultant was not sure what the best translation would be, more than one translation was recorded for future consultation.

In the next phase, a professional interpreter joined the study temporarily. Initially, she was not shown the Turkish sentences in the course books. Thus, when she watched the videos of the TiD sentences, she did not know what they were translations of. She was asked to translate these TiD sentences into Turkish orally, and her translations were recorded and transcribed. The researchers compared her translations with the original Turkish sentences and marked these pairs as (1) "complete match", (2) "semantic match but syntax is very different", and (3) "problematic". The last category refers to cases where the interpreter either didn't understand the TiD sentence or provided a translation very different from what the original Turkish sentence expressed. The interpreter and the second Deaf consultant discussed these problematic cases, and when they decided on a form they agreed on, the Deaf consultant was recorded signing the new TiD sentence. A similar procedure was carried out for the lexical entries in the TiDLaR Lexicon.

4.2 Challenges in ELAN annotations and in identifying the dependency relations

The challenges faced during the annotation of the data depended mainly on the design decisions and the nature of the corpus. As explained in Sect. 3, ELAN annotations contain dependent tiers and a dynamic lexicon. Moreover, the syntactic annotations are based on the representation of the dependency relations between constituents. Both of these properties of the corpus design forced the researchers to make certain decisions regarding annotation and linguistic analysis. Some of these are discussed below.

4.2.1 Annotating simultaneity

One commonly encountered challenge in annotating sign language utterances is related to the simultaneity property of sign languages (Johnston 2016). Non-manual markers co-occuring with manual signs and use of classifiers and buoys are among the most common examples of simultaneity and were discussed in Sect. 3. The challenge faced during the annotation of the non-manual markers is related to the hierarchical dependency of tiers of the annotation scheme adopted in this study (Sect. 3.1). When a non-lexical non-manual marker spreads over more than one

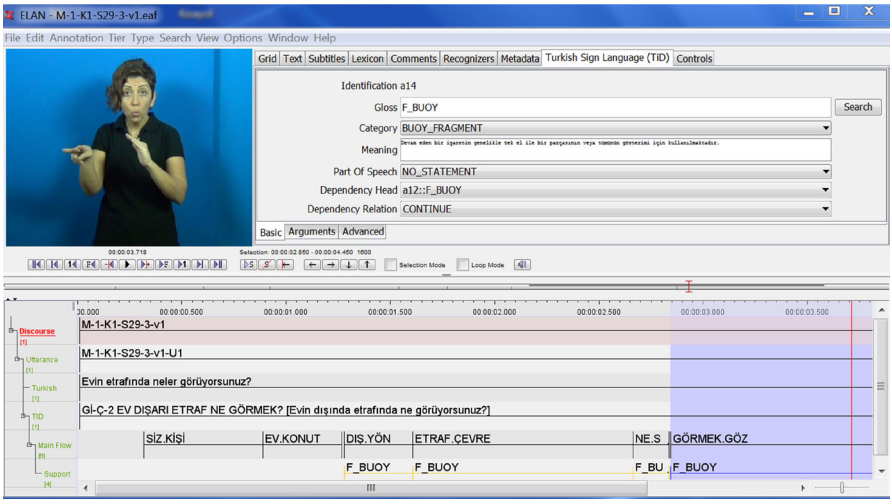

Fig. 7 Annotating simultaneity: a fragment buoy on the non-dominant hand Eng.: YOU HOUSE OUTSIDE AROUND WHAT SEE –> "What do you see outside/around the house?"

sign, this has to be indicated by entering that non-manual marker for each sign the marker spreads over.

Another challenge was annotating the relations of the manual signs that were articulated with different hands simultaneously. Most of these cases involve a classifier or a buoy on the non-dominant hand. If it is a classifier, then the sign is glossed as a classifier with information on what it represents, for instance, as PICTURE.CL (Sect. 3). If it is a buoy, then it is glossed in a way that shows its type, for instance, in the case of a fragment buoy, as F_BUOY. These glosses are entered in the SupportFlow, and to reflect the simultaneity, their time intervals are aligned with the time intervals of the signs whose glosses are in the MainFlow.

Figure 7 below exemplifies the annotation of a fragment buoy. The signer signs the lexical sign EV.KONUT 'house' and its gloss is entered in the MainFlow. As she further signs DIŞ.YÖN 'out(side)' with the dominant hand, she holds a fragment of the sign EV.KONUT on the non-dominant hand as a fragment buoy. Since the fragment buoy continues to be held until the rest of the utterance, its gloss is aligned with all of the glosses in the MainFlow. Note that since the SupportFlow tier is dependent on the MainFlow tier (see Sect. 3), a separate gloss has to be entered in the SupportFlow for each gloss in the MainFlow even though all these separate glosses refer to the same sign (in this example, the fragment buoy of the sign EV.KONUT 'house').

Many utterances in the Mathematics course book express spatial relations such as "The cat is under the dog.", "Which animal is above the rooster?" etc. The expression of these relations in TiD does not involve prepositions such as *under* and *above* with fixed pronunciations. Rather, the signing space is used to represent these spatial relations topographically, and classifiers represent the entities involved. Expressing the spatial relations may require simultaneous use of the hands and the movement of the hands. For instance, in the case of 'The cat is under the dog.', the

Table 2 Evaluation of an MT system on the developed resource (Eryiğit 2017)

	WER	PER	BLEU
# All sentences	0.39	0.23	0.47
# Simple sentences	0.37	0.24	0.52
# Compound sentences	0.43	0.22	0.37
# Short sentences ($n \leq 4$)	0.35	0.19	0.64
# Average length sentences ($4 < n \leq 8$)	0.39	0.22	0.39
# Long sentences ($8 < n$)	0.51	0.41	0.24

non-dominant hand represents the dog and is kept steady while the dominant hand represents the cat and moves downwards below the non-dominant hand. The annotation scheme enables the annotator to identify the signs as classifiers representing specific entities as, for instance, CL^CAT and their movement using the feature "Classifier_Movement" provided by the TiD add-on.

4.2.2 Identification of dependency relations

One example of TiD data that required linguistic (re)analysis and the modification of glossing for the purposes of identification of dependency relations is the use of a sign formerly glossed simply as POSS, as in POSS DRESS 'her dress'. The data contain straightforward cases such as POSS DRESS where POSS can be analyzed as a possessive pronominal determiner such as *her* and linked to DRESS via a possession relation. However, the same sign also occurs together with an overt possessor as in ANIMAL POSS FOOT 'the animal's foot'. Since the sign POSS here does not express the possessor of the foot in this example, but ANIMAL does, in this utterance POSS cannot be analyzed as a possessive pronominal determiner. To circumvent the problem, we proposed two homophonous POSS signs: POSS-PRONOUN and POSS-ADPOSITION. The latter one is analyzed as the counterpart of a preposition such as 'of' in English. Since TiD is a head-final language, this sign is analyzed as a postposition mediating the possession relation between the possessor ANIMAL and the possessee FOOT. Thus, in the dependency relation annotations, in an example such as ANIMAL POSS-ADP FOOT, ANIMAL is linked to POSS-ADPOSITION via the *CASE* relation. A similar situation holds for the annotation of negation in TiD, which is expressed at the morphological level in Turkish. A new dependency relation (*NEGATIVE*) has been added in order to express the relations emerging from the negation markers.

Another challenge was annotating holds (with classifiers or buoys) in the SupportFlow with different signs accompanying them in the MainFlow. As explained in the previous section, since the SupportFlow tier is dependent on the MainFlow tier, a continuous buoy or classifier sign co-occuring with more than one sign in the MainFlow was annotated as a separate token (Fig. 7). In terms of dependency, these separate signs were linked to their first token in the tier via the *CONTINUE* relation. The relation between a lexical item and its classifier is specified with the *COREFERENCE* relation.

5 Evaluation and discussion

This article has introduced our approach to generating the first parallel treebank between a sign language and an auditory-vocal language for use in machine translation studies. Although Turkish (being from a different language family than English and many European languages) poses interesting challenges for NLP studies, we believe the approach introduced here may be used for developing similar resources for other language pairs by making the necessary adaptations. Similar to the unification studies for POS tag and dependency representations [i.e. Universal Dependencies Project, Nivre et al. (2016)], we believe future studies on developing a universal machine readable knowledge representation scheme for sign languages will speed up machine translation studies in this field.

The resource described in this article has been recently used and evaluated within a rule-based machine translation system (Eryiğit 2017) (from Turkish to TiD). In this study, a transfer-based machine translation approach relying both on syntactic and semantic transfer has been adopted. The input to the component of translation rules is the automatic morphological and syntactic analysis of the source language (produced via a Turkish NLP pipeline) and the output which goes to feed the animation layer is the generated machine readable representation of the target language (TiD). In cases where it was not possible to find an equivalent sign entry with the same lexical sense of a Turkish input, the senses were mapped to concepts for semantic transfer adapted from the "Lexicon Model for Ontologies (LEMON), McCrae et al. (2011)". Table 2 [from Eryiğit (2017)] provides the success rates (based on WER, PER and BLEU metrics) of the above mentioned MT sytem on the developed resource. The BLEU score over all sentences was measured as 0.47 and as 0.64 on short sentences.

6 Conclusion

This article has introduced the first comprehensive machine-readable Turkish Sign Language resource (TiDLaR) built up as a result of an interdisciplinary endeavor. Scientists from linguistics, natural language processing, robotics and 3D animation fields designed TiDLaR both from a linguistic and computer science perspective hoping that it could serve as a valuable resource for both communities. For instance, linguists studying the grammar of TiD would easily search for different types of agreement verbs and different types of classifiers (cf. Appendix, Table 3 on the distribution of the parts-of-speech of the TiD data), or the distribution of 28 different non-manual markers (Table 4). Researchers working on Turkish-TiD machine-translation systems would find a prototype model both to test their MT systems and to feed the animation environments, which would obtain the required information about what should be included or suppressed over the basic motion data in order to move the avatar naturally and smoothly to produce the necessary visual sign utterance.

TiDLaR is composed of two components that interact with each other: a lexical database (available as an online interactive dictionary platform http://www.tid.itu.

edu.tr/TidGlossary/) and a parallel corpus of annotated Turkish-TiD sentence pairs, which will be available to researchers for academic purposes upon the publication of this manuscript. In addition to sign language specific annotations following the recently introduced TiD knowledge representation (Eryiğit et al. 2016), the parallel corpus contains also annotations at the morphological and syntactic (based on dependency formalism) levels, which makes it the first parallel treebank between a sign language and an auditory-vocal language.

Acknowledgements We are grateful for the support of our signers Jale Erdul, Elvan Tamyürek Özparlak, Neslihan Kurt, our Project advisors Prof. Dr. Sumru Özsoy and Hasan Dikyuva, and of our project members Pınar Uluer, Neziha Akalın, Kenan Kasarcı, Nevzat Kırgıç, Cüneyd Ancın. Finally, we want to thank our three reviewers for insightful comments and suggestions that helped us improve the final version of the article.

Appendix

See Tables 3 and 4.

Table 3 Distribution of the parts-of-speech tags in the TiD corpus

Coarse POS Tag	Fine POS Tag	Percentage
VERB	*VERB*	255 (15.2%)
	VERB_AGREEMENT_SINGLE	20 (1.2%)
	VERB_AGREEMENT_DOUBLE_FWD	20 (1.2%)
	VERB_SPATIAL	14 (0.8%)
	VERB_AGREEMENT_DOUBLE_BWD	7 (0.4%)
NOUN	*NOUN*	571 (34.0%)
	NOUN_PROPER_PERSON	13 (0.8%)
	NOUN_PROPER_ORGANIZATION	1 (0.1%)
PRONOUN	*PRONOUN*	56 (3.3%)
	PRONOUN_PERSONAL	169 (10.1%)
	PRONOUN_POSESSIVE	48 (2.9%)
POINTING BUOY	*POINTING_LOCATION*	57 (3.4%)
	BUOY_FRAGMENT	36 (2.1%)
FINGER SPELL	*FINGER_SPELL*	6 (0.4%)
CLASSIFIER	*CLASSIFIER_ENTITY_LOCATIVE*	29 (1.7%)
	CLASSIFIER_ENTITY_MOVEMENT	11 (0.7%)
	CLASSIFIER_SIZE_AND_SHAPE	18 (1.1%)
	CLASSIFIER_HANDLING	3 (0.2%)
ADJECTIVE_ADVERB	*ADJECTIVE_ADVERB*	201 (12.0%)
ADPOSITION	*ADPOSITION*	72 (4.3%)
CONJUNCTION	*CONJUNCTION*	29 (1.7%)
NUMERAL	*NUMERAL*	25 (1.5%)
PARTICLE	*PARTICLE*	16 (0.9%)

Table 4 Distribution of the non-manual markers in the TiD corpus

Non-manual name	Non-manual type	Percentage
NOISE/CHIN/TONGUE	*CHIN_DOWN*	89 (12.2%)
	CHIN_UP	15 (2.0%)
	TOUNGE_OUT	2 (0.3%)
EYE/EYEBROW	*EYEBROW_RAISING*	122 (16.7%)
	EYEBROW_FROWNING	54 (7.4%)
	EYE_NARROWED	43 (5.9%)
	EYE_EYEGAZE	36 (4.9%)
	EYE_WIDE_OPEN	31 (4.2%)
	EYE_LOOKING_DOWN	4 (0.5%)
	EYE_EYEGAZE_LEFT	3 (0.4%)
	EYE_EYEGAZE_RIGHT	3 (0.4%)
	EYE_LOOKING_UP	1 (0.1%)
HEAD	*HEAD_NOD*	109 (14.9%)
	HEAD_FORWARD	73 (10.0%)
	HEAD_TILT_RIGHT	34 (4.6%)
	HEAD_SHAKE	26 (3.6%)
	HEAD_SIDEWARD	23 (3.1%)
	HEAD_TILT_LEFT	10 (1.4%)
	HEAD_BACKWARD	3 (0.4%)
MOUTH	*MOUTH_MOUTHING*	26 (3.6%)
	MOUTH_PUFFED	2 (0.3%)
	MOUTH_LIP_PURSED	1 (0.1%)
	MOUTH_CHEEK_PUFFED	1 (0.1%)
TORSO/SHOULDER	*TORSO_RIGHT*	8 (1.1%)
	TORSO_LEFT	6 (0.8%)
	TORSO_FORWARD	5 (0.7%)
	TORSO_BACKWARD	2 (0.3%)

References

Ahrenberg, L. (2007). Lines: An English-Swedish parallel treebank. In *Proceedings of the 16th Nordic conference of computational linguistics*, pp. 270–274, Tartu.

Atalay, N. B., Oflazer, K. & Say, B. (2003). The annotation process in the Turkish treebank. In *Proceedings of the 4th international workshop on linguistically interpreteted corpora*, pp. 33–38, Budapest.

Boz, S., Özçelik, U., & Kaygusuz, Çağla. (2013). *Matematik 1* (4th ed.). Milli Eğitim Bakanlığı Yayınları, Ankara: T.C.

Bungeroth, J. & Ney, H. (2004). Statistical sign language translation. In *Proceedings of the 6th workshop on representation and processing of sign languages at the 4th international conference on language resources and evaluation*, pp. 105–108, Lisbon.

Bungeroth, J., Stein, D., Dreuw, P., Ney, H., Morrissey, S., Way, A. & Van Zijl, L. (2008). The ATIS Sign Language corpus. In *Proceedings of the 6th international conference on language resources and evaluation*, pp. 2943–2946, Marrakech.

Bungeroth, J., Stein, D., Dreuw, P., Zahedi, M. & Ney, H. (2006). A German Sign Language corpus of the domain weather report. In *Proceedings of the 5th international conference on language resources and evaluation*, pp. 2000–2003, Genoa.

Camgöz, N. C., Kindiroglu, A. A., Karabüklü, S., Kelepir, M., Özsoy, A. S. & Akarun, L. (2016). BosphorusSign: A Turkish Sign Language recognition corpus in health and finance domains. In *Proceedings of the 10th international conference on language resources and evaluation*, pp. 1383–1388, Portorož.

Cmejrek, M., Curín, J., Hajic, J. & Havelka, J. (2005). Prague Czech-English dependency treebank: resource for structure-based MT. In *Proceedings of the 11th annual conference of the European Association for Machine Translation*, pp. 73–78, Budapest.

Costello, B., Herrmann, A., Mantovan, L., Pfau, R. & Sverrisdottir, R. (2017). Section 3.10.1 Numerals. In Quer, J., Cecchetto, C., Donati, C., Geraci, C., Kelepir, M., Pfau, R. & Steinbach, M. (eds) *SignGram Blueprint: A guide to sign language grammar writing*, pp. 148–151, de Gruyter, Berlin, Boston.

Crasborn, O. & Sloetjes, H. (2008). Enhanced ELAN functionality for sign language corpora. In *Proceedings of the 3rd workshop on the representation and processing of sign languages: Construction and exploitation of sign language corpora at the 6th international conference on language resources and evaluation*, pp. 39–43, Marrakech.

Crasborn, O. A. & Zwitserlood, I. (2008). The corpus NGT: An online corpus for professionals and laymen. In *Proceedings of the 3rd workshop on the representation and processing of sign languages: Construction and exploitation of sign language corpora at the 6th international conference on language resources and evaluation*, pp. 44–49.

Crasborn, O., Bank, R., Zwitserlood, I., van der Kooij, E., de Meijer, A., & Safar, A. (2015). *Annotation conventions for the corpus NGT*. Ms: Radboud University Nijmegen.

Cuřín, J., Čmejrek, M., Havelka, J. & Kuboň, V. (2004). Building a parallel bilingual syntactically annotated corpus. In *Proceedings of the international conference on natural language processing*, pp. 168–176, Hyderabad.

Dalkılıç, H., & Gölge, N. (2013). *Hayat Bilgisi 1* (4th ed.). Milli Eğitim Bakanlığı Yayınları, Ankara: T.C.

De Vos, C., van Zuilen, M., Crasborn, O. & Levinson, S. (2015). NGT interactive corpus. *MPI for psycholinguistics, the language archive*, https://hdl.handle.net/1839/00-0000-0000-0021-8357-B@view.

Demiroğlu, R., & Gökahmetoğlu, E. (2013). *Türkçe 1* (4th ed.). Milli Eğitim Bakanlığı Yayınları, Ankara: T.C.

DeNeefe, S., Knight, K., Wang, W. & Marcu, D. (2007). What can syntax-based MT learn from phrase-based MT? In *Proceedings of the joint conference on empirical methods in natural language processing and computational natural language learning*, pp. 755–763, Prague.

Dikyuva, H., Makaroğlu, B., & Arık, E. (2017). *Turkish sign language grammar*. Ankara: Ministry of Family and Social Policies Press.

Eryiğit, G. (2007a). ITU validation set for Metu-Sabancı Turkish treebank.

Eryiğit, G. (2007b). ITU treebank annotation tool. In *Proceedings of the linguistic annotation workshop at the 40th annual meeting on association for computational linguistics*, pp. 117–120, Prague.

Eryiğit, G. (2014). ITU Turkish NLP web service. In *Proceedings of the demonstrations at the 14th conference of the European chapter of the association for computational linguistics*, pp. 1–4, Gothenburg.

Eryiğit, C. (2017). Text to sign language machine translation system for Turkish. Ph.D. thesis, Istanbul Technical University, Istanbul.

Eryiğit, G., Adalı, K., Torunoğlu-Selamet, D., Sulubacak, U. & Pamay, T. (2015). Annotation and extraction of multiword expressions in Turkish treebanks. In *Proceedings of the human language technology conference at the North American Chapter of the Association for Computational Linguistics*, pp. 70–76, Denver, CO.

Eryiğit, C., Köse, H., Kelepir, M., & Eryiğit, G. (2016). Building machine-readable knowledge representations for Turkish Sign Language generation. *Knowledge-Based Systems*, *108*, 179–194.

Galley, M., Graehl, J., Knight, K., Marcu, D., DeNeefe, S., Wang, W. & Thayer, I. (2006). Scalable inference and training of context-rich syntactic translation models. In *Proceedings of the 21st international conference on computational linguistics and the 44th annual meeting of the association for computational linguistics*, pp. 961–968, Sydney.

Galley, M., Hopkins, M., Knight, K. & Marcu, D. (2004). What's in a translation rule? In *Proceedings of the human language technology conference at the North American Chapter of the Association for Computational Linguistics*, pp. 273–280, Boston, MA.

Hanke, T. & Storz, J. (2008). iLex–a database tool for integrating sign language corpus linguistics and sign language lexicography. In *Proceedings of the 3rd workshop on the representation and processing of sign languages at the 6th international conference on language resources and evaluation*, pp. 64–67, Marrakech.

Johnston, T. (2008). Corpus linguistics and signed languages: No lemmata, no corpus. In *The 3rd workshop on the representation and processing of sign languages: Construction and exploitation of sign language corpora at the 6th international conference on language resources and evaluation*, Marrakech.

Johnston, T. (2016). Auslan corpus annotation guidelines. *Centre for Language Sciences, Department of Linguistics, Macquarie University (Sydney) and La Trobe University (Melbourne)*, http://media.auslan.org.au/attachments/Auslan_Corpus_Annotation_Guidelines_November2016.pdf.

Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., et al. (2007). Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th annual meeting of the association for computational linguistics: Companion volume—proceedings of the demo and poster sessions*, pp. 177–180, Prague.

Koizumi, A., Sagawa, H. & Takeuchi, M. (2002). An annotated Japanese Sign Language corpus. In *Proceedings of the 3rd international conference on language resources and evaluation*, pp. 927–930, Las Palmas.

Kubuş, O. (2008). An analysis of Turkish Sign Language (TiD) phonology and morphology. Master's thesis, Middle East Technical University, Ankara.

Leeson, L., Saeed, J., Leonard, C., Macduff, A. & Byrne-Dunne, D. (2006). Moving heads and moving hands: Developing a digital corpus of Irish Sign Language: The 'Signs of Ireland' corpus development project. In *Proceedings of the information technology and telecommunications conference*, Carlow.

Liddell, S. K. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge: Cambridge University Press.

McCrae, J., Spohr, D. & Cimiano, P. (2011). Linking lexical resources and ontologies on the semantic web with Lemon. In *The semantic web: Research and applications*, pp. 245–259, Berlin, Springer.

Megyesi, B., Dahlqvist, B., Pettersson, E. & Nivre, J. (2008). Swedish—Turkish parallel treebank. In *Proceedings of the 6th international conference on language resources and evaluation*, pp. 470–473, Marrakech.

Miller, C. (2001). Section I: Some reflections on the need for a common sign notation. *Sign Language & Linguistics, 4*(1), 11–28.

Neidle, C., Sclaroff, S., & Athitsos, V. (2001). Signstream: A tool for linguistic and computer vision research on visual-gestural language data. *Behavior Research Methods, Instruments, & Computers*, *33*(3), 311–320. https://doi.org/10.3758/BF03195384.

Nivre, J., de Marneffe, M.-C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., McDonald, R., Petrov, S., Pyysalo, S., Silveira, N., et al. (2016). Universal dependencies v1: A multilingual treebank collection. In *Proceedings of the 10th international conference on language resources and evaluation*, pp. 1659–1666, Portorož.

Oflazer, K., Say, B., Hakkani-Tür, D. Z., & Tür, G. (2003). Building a Turkish treebank. In A. Abeillé (Ed.), *Treebanks: Building and using parsed corpora* (pp. 261–277). London: Kluwer.

Östling, R., Börstell, C., Gaärdenfors, M. & Wirén, M. (2017). Universal dependencies for Swedish Sign Language. In *Proceedings of the 21st Nordic conference on computational linguistics*, pp. 303–308, Gothenburg.

Othman, A., Tmar, Z. & Jemni, M. (2012). Toward developing a very big sign language parallel corpus. In *Proceedings of the international conference on computers for handicapped persons*, pp. 192–199, Paris.

Özsoy, S., Arık, E., Göksel, A., Kelepir, M. & Nuhbalaoğlu, D. (2013). Documenting Turkish sign language: A report on a research project. In *Current directions in TiD research*, pp. 55–70, Cambridge Scholars.

Pamay, T., Sulubacak, U., Torunoğlu-Selamet, D. & Eryiğit, G. (2015). The annotation process of the ITU web treebank. In *Proceedings of the 9th linguistic annotation workshop at the North American Chapter of the Association for computational linguistics*, pp. 95–101, Denver, CO.

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, *227*(1), 1–15.

Pfau, R. & Quer, J. (2010). Nonmanuals: Their prosodic and grammatical roles. In *Sign languages*, pp. 381–402. Cambridge, Cambridge University Press.

Prillwitz, S., Hanke, T., König, S., Konrad, R., Langer, G. & Schwarz, A. (2008). DGS corpus project–development of a corpus based electronic dictionary German Sign Language/German. In *Proceedings of the 3rd workshop on the representation and processing of sign languages at the 6th international conference on language resources and evaluation*, pp. 159–164, Marrakech.

Şahin, M., Sulubacak, U. & Eryiğit, G. (2013) Redefinition of Turkish morphology using flag diacritics. In *Proceedings of the 10th symposium on natural language processing*, Phuket.

Schembri, A., Fenlon, J., Rentelis, R., Reynolds, S., & Cormier, K. (2013). Building the British Sign Language corpus. *Language Documentation & Conservation*, *7*, 136–154.

Selçuk-Şimşek, M., & Çiçekli, I. (2017). Bidirectional machine translation between Turkish and Turkish Sign Language: A data-driven approach. *International Journal on Natural Language Computing*, *6*(3), 33–46.

Steinbach, M. (2012). Plurality. In *Sign language: An international handbook*, pp. 112–136, De Gruyter Mouton.

Sulubacak, U. & Eryiğit, G. (2013). Representation of morphosyntactic units and coordination structures in the Turkish dependency treebank. In *Proceedings of the 4th workshop on statistical parsing of morphologically rich languages at the conference on empirical methods on natural language processing*, p. 129, Seattle, WA.

Sulubacak, U., Gokirmak, M., Tyers, F., Çöltekin, Ç., Nivre, J. & Eryiğit, G. (2016a). Universal dependencies for Turkish. In *Proceedings of the 26th international conference on computational linguistics*, pp. 3444–3454, Osaka.

Sulubacak, U., Pamay, T. & Eryiğit, G. (2016b). IMST: A revisited Turkish dependency treebank. In *Proceedings of the 1st international conference on turkic computational linguistics at the international conference on computational linguistics and intelligent text processing*, pp. 1–6, Konya.

Sulubacak, U., & Eryigit, G. (2018). Implementing universal dependency, morphology, and multiword expression annotation standards for Turkish language processing. *Turkish Journal of Electrical Engineering & Computer Sciences*, *26*(3), 1662–1672.

Su, H.-Y., & Wu, C.-H. (2009). Improving structural statistical machine translation for sign language with small corpus using thematic role templates as translation memory. *IEEE Transactions on Audio, Speech, and Language Processing*, *17*(7), 1305–1315.

Swedish Sign Language Corpus Project, U. D. (2017). Universal dependencies for Swedish Sign Language. *Stockholm University*, https://www.ling.su.se/english/research/research-projects/sign-language/swedish-sign-language-corpus-project-1.59270.

Tesnière, L. (1959). *Eléments de syntaxe structurale*. Klincksieck: Librairie C.

Tinsley, J., Hearne, M. & Way, A. (2009). Exploiting parallel treebanks to improve phrase-based statistical machine translation. In *Proceedings of the international conference on intelligent text processing and computational linguistics*, pp. 318–331, Mexico City.

Uchimoto, K., Zhang, Y., Sudo, K., Murata, M., Sekine, S. & Isahara, H. (2004). Multilingual aligned parallel treebank corpus reflecting contextual information and its applications. In *Proceedings of the workshop on multilingual linguistic resources at the 42th annual meeting on association of computational linguistics*, pp. 63–70, Barcelona.

Wallin, L. & Mesch, J. (2015). Swedish sign language corpus. In *Proceedings of digging into signs workshop: Developing annotation standards for sign language corpora*, London.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A. & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *Proceedings of the 5th international conference on language resources and evaluation*, pp. 1556–1559, Genoa.

Zwitserlood, I. (2012). Classifiers. In *Sign languages: An international handbook*, pp. 158–186, Mouton de Gruyter.

Zwitserlood, I., Perniss, P., & Özyürek, A. (2012). An empirical investigation of expression of multiple entities in Turkish Sign Language (TİD): Considering the effects of modality. *Lingua*, *122*(14), 1636–1667.