Towards parsing language learner utterances in context

Christine Köhn and Wolfgang Menzel
Department of Informatics
Universität Hamburg
{ckoehn, menzel}@informatik.uni-hamburg.de

Abstract

Intelligent computer-assisted language learning (ICALL) systems give feedback on grammatical errors based on the syntactic analysis of an utterance. However, parsing a language learner sentence in isolation can result in multiple interpretations, which even humans may not be able to resolve. For deriving a helpful error diagnosis, an interpretation that is consistent with the intended meaning is needed. To achieve this, we propose an approach that models information about the context of a sentence by means of conceptual relationships and integrates them into the decision procedure of a syntactic-semantic parser. This way, we obtain a syntactic structure which complies best with a given context model. We identify and categorize cases where the context influences both the structure and the error diagnosis derived from it.

1 Introduction

For diagnosing grammatical errors in a language learner sentence, its syntactic structure is needed. Based on the syntactic structure, an error diagnosis and a correction can be inferred. A syntactic analysis based on the sentence alone is not sufficient for this purpose, because the context in which a language learner utters a sentence sometimes plays a crucial role for grammatical error diagnosis. Depending on the context, a sentence can be judged as correct or erroneous. Furthermore, among the different error diagnoses for a faulty utterance, the context can help to choose the ones which are more likely than the others.

Different types of context information can be exploited to interpret a language learner utterance such as the learner's native language and the exercise. In this paper, we focus on a subset of the context, namely the situation the learner is writing about. In the following, examples will be given to illustrate different cases where the context of an utterance influences the error diagnoses.

The German[1] sentence in Example 1 is well-formed. However, *See* is ambiguous with respect to grammatical gender and if *See* is meant to refer to a lake instead of the sea, the grammatical gender of *See* is masculine. Therefore, the corresponding article needs to be masculine as well and the sentence is erroneous. Example 2 shows the corrected sentence.

E1	Das	Haus	liegt	an	**der**$_{\text{f,dat}}$	See$_{\text{f,dat}}$
	The	house	is located	by	the	sea
E2	Das	Haus	liegt	an	**dem**$_{\text{m,dat}}$	See$_{\text{m,dat}}$
	The	house	is located	by	the	lake

Assuming that *See* denotes a lake in Example 1, two probable error diagnoses are available: The student inflected the article correctly (dative case) but assumed the wrong grammatical gender for the word *See* (*der*: feminine, *dem*: masculine), or the student chose the correct gender (masculine) but did not use the right case (*der*: nominative, *dem*: dative). Which of these error diagnoses expresses the student's misconception best cannot be determined without further background knowledge but could be achieved through student modeling. This, however, is not considered in this paper.

A clearly erroneous sentence is shown in Example 3 where the article *die* needs to be replaced. Depending on whether *See* denotes the sea or a lake, it can be corrected into Example 1 or Example 2, respectively. If *See* means the sea, the student probably selected the appropriate gender (feminine) but the wrong case (*die*: nominative or accusative, *der*: dative). If *See* means a lake, not only the case is wrong (same as before) but also the gender (*die*: feminine, *dem*: masculine).

E3	*Das	Haus	liegt	an	**die**$_{\text{f,nom/acc}}$	See$_{\text{f/m,dat}}$
	The	house	is located	by	the	sea/lake

[1] All of the examples given in this paper are in German.

Proceedings of the Int. Conference of the German Society for Computational Linguistics and Language Technology,
pages 144–153, University of Duisburg-Essen, Germany, Sep 30–Oct 2 2015.

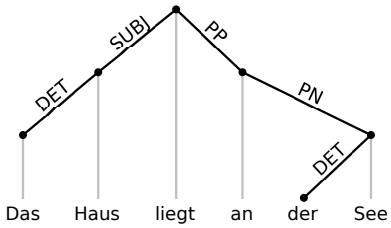
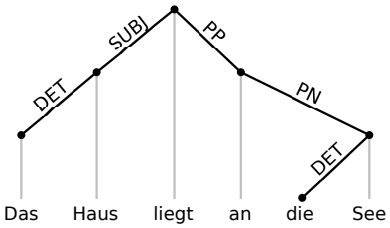
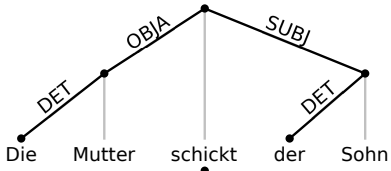
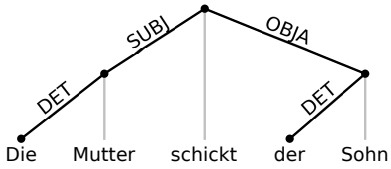
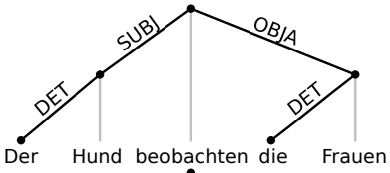
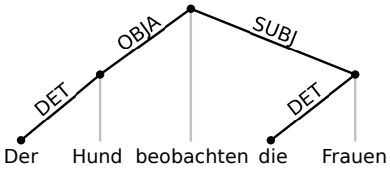
Synt. structure	Distinctions induced by the context: Erroneous vs. correct	Distinctions induced by the context: Different diagnoses and corrections
invariant to context	 Example 1	 Example 3
dependent on context	  Example 4	  Example 6

Table 1: Overview of the example sentences **E1**, **E3**, **E4** and **E6** illustrating the influence of the utterance context on syntactic parsing and error diagnosis (see text). Labels: `DET` (determiner of a noun), `SUBJ` (subject of a verb), `OBJA` (accusative object) and `PP`/`PN` (prepositional phrase/complement)

These examples illustrate that the context can help to distinguish between erroneous and correct sentences as well as narrow down the set of error diagnoses. The syntactic structures for the above examples are shown in Table 1.[2] For examples 1 to 3, the syntactic structures are the same, regardless of the sea/lake distinction. In other cases, the syntactic structure of a sentence might depend on the context: Example 4 shows a sentence with a structural ambiguity. It is correct but only as long as we assume that the mother is the object of the sentence and the son is the subject.

E4 (*)Die$_{\text{nom/}\mathbf{acc}}$ Mutter$_{\text{nom/}\mathbf{acc}}$ schickt **der**$_{\text{nom}}$ Sohn$_{\text{nom}}$
The mother is sending the son
(The son is sending the mother)

If the sentence in Example 4 is uttered in a context where the mother is sending the son, then the mother is the subject and the son the object (see Table 1 for the syntactic structures). Under this reading, the sentence becomes erroneous: Being the object, *der Sohn* has to be accusative case but it is nominative case. Therefore, the article needs to be changed to *den* (Example 5).

E5 Die$_{\text{nom}}$ Mutter$_{\text{nom}}$ schickt **den**$_{\text{acc}}$ Sohn$_{\text{acc}}$
The mother is sending the son

Cases where the sentence is obviously erroneous are similarly difficult. Then, different syntactic interpretations for that sentence might be possible, and, consequently, the error diagnoses inferred from them may differ. Example 6 shows such a case: A syntactic structure which assigns dog as the subject and (several) women as the object implies a verb form error because the subject *Hund* and the verb *beobachten* do not agree in number (*Hund*: singular, *beobachten*: plural). To correct the error, the verb has to be changed to singular (Example 7). Given a context in which the women are watching a dog, another syntactic structure which assigns dog as the object and women as the subject (see Table 1) is more adequate. Based on this interpretation, a different error diagnosis can be obtained: The object *der Hund* has the wrong case (nominative instead of accusative). Thus, the article *der* needs to be corrected to *den* (Example 8).

[2]The syntactic structures are displayed as dependency trees. We apply the annotation scheme by Foth (2006). A short description of the labels in English can be found in Foth et al. (2014).

E6 *Der$_{\text{sg,nom}}$ Hund$_{\text{sg,nom}}$ beobachten$_{\text{pl}}$ die$_{\text{pl,nom/acc}}$ Frauen$_{\text{pl,nom/acc}}$
The dog are watching the women

E7 Der Hund **beobachtet**$_{\text{sg}}$ die Frauen
The dog is watching the women

E8 **Den**$_{\text{acc}}$ Hund$_{\text{acc}}$ beobachten die Frauen
The dog are watching the women
(The women are watching the dog)

The examples show that the context of an utterance influences its grammaticality judgment as well as the error diagnoses. This paper focuses on situations where the context gives rise to different syntactic structures and to different error diagnoses, which can be derived from them. We will adopt a method for modeling the context information and integrating it into the decision procedure of a syntactic parser to obtain the syntactic structure that is most plausible with respect to the given context.

The remainder of this paper is structured as follows: The parsing formalism (Section 3) and the context models (Section 4) we work with will be explained. Section 5 will illustrate, by means of an example, an approach for integrating context information into parsing. Section 6 will describe one core mechanism of the context integration process, the selection of referents. In Section 7, we will systematically identify and categorize cases where context integration could help to distinguish different syntactic structures. Limitations of this approach will be discussed in Section 8. Section 9 will conclude the paper.

2 Related Work

When parsing language learner utterances, the parser has to be error-tolerant, and in addition, its output should contain information that can be used for generating an error diagnosis. Several approaches exist for parsing language learner sentences in order to diagnose grammatical errors, e. g., Heift (2003), Reuer (2003), Bender et al. (2004), Fortmann and Forst (2004) and Boyd (2012).

To our knowledge, there are only two approaches which integrate context information into the parsing procedure itself when analyzing language learner sentences: An early experiment was conducted by Menzel and Schröder (1999) where domain knowledge was integrated into syntactic parsing of artificially distorted sentences originating from a single sentence. More recently, Antonsen et al. (2009) have used a parser which implements the Constraint Grammar framework (Karlsson et al., 1995) for analyzing learner sentences in an ICALL system where the user answers questions asked by the system. The parser uses two rule sets: The first set disambiguates the input partially in order to find appropriate readings for erroneous input. The second set contains rules that flag errors in the input. Context information such as verb tense and case of the interrogative determiner from the question restricts the interpretation of the user's answer. In contrast to Antonsen et al. (2009), we do not expect the input to be questions and answers but free-form text from writing exercises such as picture description tasks where a question is not necessarily available for the interpretation of the input.

Some ICALL applications extract a semantic interpretation from the syntactic structure of a sentence for further processing but the semantic information is not used for parsing itself. Hahn and Meurers (2012), for example, evaluate the meaning of short answers. For this purpose, MaltParser (Nivre et al., 2007), a dependency parser, obtains a syntactic structure from which a semantic representation is derived.

3 Parsing with constraint relaxation

A dependency parser constructs a structural description, the dependency tree, of an input sentence by assigning each word (the dependent) to a regent (either another word or the special root node). The parser labels these dependencies to characterize the relationship between dependent and regent. For parsing, the Weighted Constraint Dependency Grammar (WCDG) parser is used (Foth et al., 2004). In the WCDG formalism, well-formedness conditions for dependency trees are expressed as constraints. Each constraint stipulates a condition which a set of edges should satisfy. In addition, it is graded with a weight from $[0,1]$, which indicates how severe it is if a dependency tree violates the particular constraint: The closer the weight is to 0, the more severe the violation. For example, having two subjects as dependents of a verb is more severe than assigning a subject to a verb which does not agree with it. The WCDG parser finds the structural description that best adheres to all constraints defined in the grammar: The best structure s is defined as

$$s = \arg\max_{s'} \prod_{c \in \text{Constraints}} \text{weight}(c)^{\text{n}(c,s')} \quad (1)$$

where weight(c) is the weight of constraint c and $\mathrm{n}(c, s')$ is the number of times c is violated in the structure s', i. e. the number of times the parser had to relax c to obtain s'. The parser performs a heuristically-guided transformation-based search, which starts with an initial structure and tries to resolve constraint violations iteratively.

The constraint relaxation mechanism makes the parser robust to ill-formed input (Foth et al., 2005). As a byproduct, the unresolved constraint violations can be used to infer error diagnoses for the input sentence.

WCDG achieves a labeled attachment accuracy on the learner corpus CREG-109 (Ott and Ziai, 2010) of 79.28% (Krivanek and Meurers, 2011). On newspaper data, an accuracy of 81.42%[3] (Krivanek and Meurers, 2011) was measured on the TüBa-D/Z corpus (Telljohann et al., 2004) and 91.0% (Foth and Menzel, 2006) on the NEGRA corpus (Brants et al., 1999). Currently, the German grammar contains more than 1000 handwritten constraints.

In the WCDG formalism, a dependency structure may span multiple levels of analysis. In addition to the syntactic level, we will use a semantic level, which contains a semantic analysis in the form of thematic roles. This level serves as an interface for the context integration.

4 Modeling context

A context model specifies high-level information about the situation, in which a sentence was uttered. In this paper, we follow McCrae (2010) and model only part of the context, i. e., actions and their participants and relationships between them. Later on, we will show that this part is sufficient to influence the syntactic analysis of possibly erroneous utterances.

A context model consists of two parts: situation-invariant knowledge (the T-Box) and situation-dependent knowledge (the A-Box), which we specify using the description logic OWL (Web Ontology Language, Bechhofer et al. (2004)).

4.1 T-Box

The T-Box (terminological box) is an ontology, which defines classes and relationships between them, e. g., MOTHER $\xrightarrow{\text{is\_a}}$ WOMAN. The IS-A relations form a hierarchical taxonomy. Each class has one or more lexicalisations, which specify the words and phrases that can be used to denote individuals of that class. Every individual described in the situation-dependent knowledge instantiates one of the concepts from the T-Box.

The T-Box can be manually created for the domain as was done in McCrae (2010) but a readily-available ontology such as GermaNet (Hamp and Feldweg, 1997) could also be exploited for that purpose.

4.2 A-Box

The A-Box (assertional box) contains individuals, which are instances of concepts in the T-Box. The A-Box defines relationships between individuals or assigns properties to them. An individual represents an object, an action or a participant who engages in an action. McCrae (2010) defines several relations which can hold between an action and its participants in terms of the thematic role the participant fulfills in that action. In Figure 1b, for example, WOMAN_1 is modeled as the AGENT of the sending action.

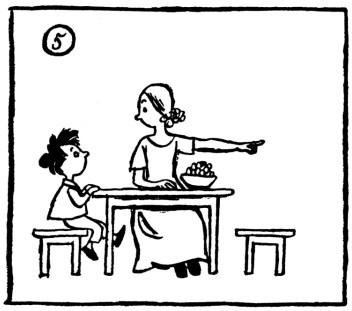
The A-Box can either be manually created as in McCrae (2010) or semi-automatically using the parser again: Simple sentences particularly authored for that purpose can be analyzed in order to extract individuals and the relationships between them automatically.

McCrae (2010) and Baumgärtner et al. (2012) use the A-Box to describe a visual scene and therefore limit the context model to include only visually perceivable information or indirectly derived information inferred from the visual input. Other scenarios are also conceivable, where the context information is derived from a textual description accompanying an exercise or from the preceding discourse.

5 Obtaining a context-induced syntactic structure

Guided by Example 4, this section will explain how the parser obtains a syntactic structure that is compatible with a given utterance context. Parsing the sentence using solely linguistic features results in a syntactic structure where the mother is the object and the son is the subject (first analysis for Example 4 in Table 1) because a parser usually prefers a structure which is as well-formed as possible. Thus, judged by this structure, the sentence has to be considered correct German. However, if the sentence

[3] The low accuracy might be caused by mismatches in the annotation scheme.



(a)

WOMAN_1	$\xrightarrow{\text{is\_instance\_of}}$	WOMAN
BOY_1	$\xrightarrow{\text{is\_instance\_of}}$	BOY
SEND.SB_1	$\xrightarrow{\text{is\_instance\_of}}$	SEND.SB
WOMAN_1	$\xrightarrow{\text{is\_AGENT\_for}}$	SEND.SB_1
BOY_1	$\xrightarrow{\text{is\_THEME\_for}}$	SEND.SB_1

(b)

Figure 1: An image from a picture story (Ohser, 1993) and the A-Box of its context model

"Die" *The*	$\xrightarrow{\text{is\_conceptualised\_by}}$	{}
"Mutter" *mother*	$\xrightarrow{\text{is\_conceptualised\_by}}$	{MOTHER}
"schickt" *is sending*	$\xrightarrow{\text{is\_conceptualised\_by}}$	{SEND.SB}
"der" *the*	$\xrightarrow{\text{is\_conceptualised\_by}}$	{}
"Sohn" *son*	$\xrightarrow{\text{is\_conceptualised\_by}}$	{SON}

(a) Words mapped to concepts

"Die"	$\xrightarrow{\text{matches}}$	{}
"Mutter"	$\xrightarrow{\text{matches}}$	{WOMAN_1}
"schickt"	$\xrightarrow{\text{matches}}$	{SEND.SB_1}
"der"	$\xrightarrow{\text{matches}}$	{}
"Sohn"	$\xrightarrow{\text{matches}}$	{BOY_1}

(b) Candidate referents for words

Figure 2: Finding referents for words in the context model

was uttered while describing a situation where actually the mother is sending the son, it is erroneous. Interchanging the subject and object in the syntactic structure so that the structure complies with the described situation would reveal that the wrong case was chosen for *der Sohn*. Therefore, we want the parser to output a context-induced syntactic structure where the mother (subject) is sending the son (object).

To deal with this problem, the parser has to be enabled to choose one syntactic structure in one context and another in a different context for the same sentence, even though the sentence might contain errors in one of these interpretations. For this purpose, we adopt a model developed by McCrae (2009) and Baumgärtner et al. (2012): By mapping propositions about the context to syntactic relationships in the sentence, the parser is guided towards a syntactic structure which conforms to a given visual context. So far, this model has only been applied to syntactically ambiguous well-formed German sentences with the objective of modulating the syntactic structure with context information, e. g., the attachment of prepositional phrases. Our goal is to analyze learner utterances, whereas the goal of the aforementioned work was to disambiguate correct German sentences. Nonetheless, their model is also capable of resolving some of the ambiguities caused by errors. When we apply that model, we are able to obtain a context-induced syntactic interpretation of the example sentence.

The sentence in Example 4 is the beginning of a genuine language learner sentence written by a learner when asked to describe the content of the picture in Figure 1 as part of a picture story. We assume that the parser receives a context model for the picture as input, i. e., a high-level description of the situation depicted (Figure 1): A woman (WOMAN_1) and a boy (BOY_1) are engaging in a sending action (SEND.SB_1) where the woman is the agent of that action (the one that is sending somebody) and the boy is the theme (the one that is being sent). Each of the individuals (WOMAN_1, BOY_1, SEND.SB_1) are instances of classes of the ontology in the T-Box.

First, a connection between the sentence and the context model is established. For each word, a matching individual (the word's referent) is searched in the A-Box: The word's lemma is compared to the lexicalisations of each class in the ontology, obtaining a set of activated concepts for each word (Figure 2a). The set of candidate referents for each word is built by adding each individual whose class is close to a concept activated by the word (Figure 2b). As a result, not only direct matches (SEND.SB_1 for "schickt") but also conceptually related individuals (WOMAN_1 for *Mutter* and BOY_1 for *Sohn*) are selected as candidate referents. Among all possible assignments of candidates to words, a scoring mechanism deter-

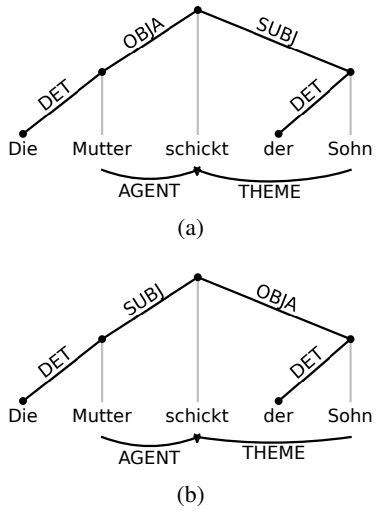


Figure 3: Analyses of Example 4

mines the best assignment. For our example, this would result in the mapping given in Figure 2b.

The information from the context model is integrated into parsing by means of constraints that access the information made available by the linking of words to their referents. One constraint could, for example, require that an agent (or theme) relation between two words in the analysis must also exist between their referents in the A-Box. By applying such constraints to the semantic level of the analysis, the attachments on this level can be pushed to reflect the relationships the word's referents engage in the A-Box. In our example, the *Mutter* would be attached to *schickt* by an `AGENT` edge, and *Sohn* would be attached as the `THEME` (Figure 3a).

If the parser's grammar contains constraints that each describe a condition either for the syntactic level or the semantic level, the parser would output the structure in Figure 3a: It complies best with the constraints on the syntactic level and on the semantic level in isolation, but the syntactic and semantic analysis contradict each other. To obtain the desired structure where *Mutter* is the subject and *Sohn* is the object, constraints which mediate between the syntactic and the semantic level need to be employed to influence the syntactic structure. Interchanging subject and object in Figure 3a can be achieved by adding the following constraints: The `THEME` should be the accusative object `OBJA` and the `AGENT` the subject `SUBJ` (in an active-voice sentence). If the weights of these constraints penalize their violation more heavily than the ones on the syntactic level, the parser obtains the structure in Figure 3b: The syntactic level is not well-formed anymore but complies with the context model. Based on this structure, a case error can be diagnosed: *der Sohn* (nominative) has to be changed to *den Sohn* (accusative).

6 Selecting Referents

Candidate referents for a word are found by comparing the lexicalisations of the concepts in the T-Box to the lemma of the word. Every individual in the A-Box which is an instance of a matching concept or an instance of a concept close to a matching concept is a candidate referent for the word. To select adequate referents for the words in the sentence, each possible assignment of referents to words is rated and the assignment with the highest score is selected.

Several measurements contribute to the overall score of an assignment (Baumgärtner et al., 2012):

Distance between concepts The suitability of a referent for a word depends on the distance of the concept instantiated by the referent to the concept activated by the word in the taxonomy. The closer the concepts, the higher the rating for assigning a referent to a word.

Incorporating the conceptual distance makes the matching process more robust because it allows to match concepts which are on different levels in the taxonomy (e. g., MOTHER and WOMAN) but even if there is no super- or subclass relationship between the two concepts it may be reasonable to establish a match if the concepts are close enough (e. g., CUP and MUG).

Similarity between word and lexicalisation The more similar the lemma of the word and the lexicalisation of the referent are on the character level, the better the assignment is rated. This makes the matching robust to slight spelling and typing errors.

Connections between referents A referent is rated higher if it is related to an individual in the A-Box which has been selected as a referent for another word.

This increases the likelihood that the appropriate referents are chosen in cases where the conceptual distance and the character similarity do not distinguish between two candidate referents. For example: If there were another woman, WOMAN_2, in the A-Box of Figure 1b, WOMAN_1 would be rated higher than WOMAN_2 in a sentence where

	SUBJ	OBJA	OBJD	GMOD	S
OBJA	•				–
OBJD	•	•			–
DET	•	•	•		–
GMOD	•	•	•		–
APP	•	•	•	•	–
REL	–	–	–	–	•

Table 2: Confusion matrix of syntactic functions. "•": Confusion is possible. "–": No example has been found. Please note that the table is symmetric: The entry for (DET, OBJA), e. g., can be found in cell (OBJA, DET). Labels: APP (apposition), REL (relative clause), S (sentence); for the other labels see Table 3.

Information in context model	Promoted syntactic attachment
AGENT	SUBJ (subject of a verb)
THEME	DET (determiner of a noun) OBJA (accusative object of a verb) SUBJ (subject of a verb)
RECIPIENT	OBJD (dative object of a verb)
OWNER	DET (determiner of a noun) GMOD (genitive modifier of a noun)

Table 3: Syntactic attachments which are promoted by the relations in the context model

SEND.SB_1 has already been selected as a referent for another word.

Influence from the syntactic structure Relationships expressed in the syntactic structure of a sentence may also be present in the A-Box. For example: If the syntactic structure assigns an attribute (e. g. "big") to a noun (e. g. "dog"), an individual which exhibits the same property is a more likely referent for that word than another individual (a dog who is big as opposed to any other dog).

The selection process does not require that all individuals from an A-Box have to be referents for words, and vice versa, not every word has to have a referent in the A-Box. Additionally, the context model does not restrict the word order and allows for a variety of word choices, since referents can be matched via the taxonomy in the T-Box. As a result, an A-Box is more general than a sentence with the textual description of the A-Box: One A-Box can be matched to numerous sentences.

The selection of referents is an iterative process, since the referents and the parser's analysis mutually influence each other. After the selection, the referents for each word are fed back to the parser, which reanalyzes the sentence. Whenever the parser finds an analysis with a higher score with respect to Equation (1), the selection of referents is renewed.

7 Structurally Ambiguous Cases

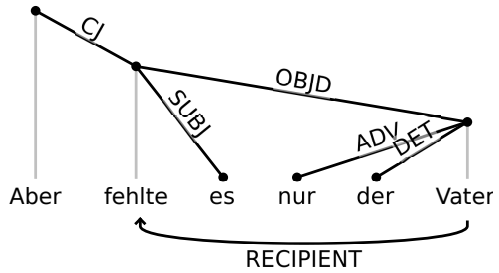
If errors are present in a sentence, the syntactic structure might be ambiguous in a way that the ambiguity can only be resolved by integrating context information into syntactic parsing (cf. Example 4 and 6 in Section 1). We systematically analyzed common language learner error types such as case and gender selection errors (Rogers, 1984) as to when they cause ambiguities with respect to syntactic functions.

We have grouped the error-induced ambiguities into confusion classes: Table 2 shows the syntactic functions (denoted by the name of the label) that could be confused with each other if no biasing context information is available. Confusions are not limited to the exchange of syntactic functions, e. g. the subject SUBJ against the accusative object OBJA, but can also affect the attachment of a word: Confusing the dative object OBJD and the genitive modifier GMOD, e. g., implies that the word's regents differ, since dative objects are attached to a verb, and genitive modifiers to a noun.

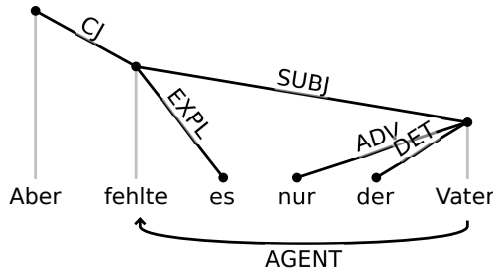
For each confusion of syntactic functions, we have identified disambiguating context information in the form of thematic roles. Four of the six thematic roles employed by McCrae (2010) are relevant for disambiguating the structural ambiguities from Table 2: AGENT, THEME, RECIPIENT and OWNER. Each role influences the attachment of words in the sentence (by means of the mediating semantic level of the analysis) in a different way, e. g., an AGENT of an action is more likely to be the subject of the sentence in active voice, whereas the RECIPIENT is more likely to be the dative object. Table 3 gives an overview of the syntactic attachments which are promoted by the relationships in the A-Box[4].

Figure 4 shows two analyses of an erroneous real-world learner sentence, which exhibits the

[4]Example sentences which exhibit the confusions of Table 2 and show the disambiguating capabilities of the thematic roles from Table 3 can be found in Köhn and Menzel (2015), the extended version of this paper.



(a) Interpreting "father" as dative object implies a word order and a case error with respect to the target hypothesis "Aber **es fehlte** nur **dem** Vater" (But only the father missed it).



(b) Interpreting "father" as the subject implies a word order error with respect to the target hypothesis "Aber **es fehlte** nur der Vater" (But only the father was missing).

Figure 4: Two different analyses of the same learner sentence (Literally: But missed it only the father).

SUBJ/OBJD confusion. Depending on the context, the father is either the subject or the object. Without any context information, both WCDG and the machine learning-based TurboParser[5] (Martins et al., 2009; Martins et al., 2013) obtain the analysis in Figure 4a (without the RECIPIENT edge). However, the learner wrote about a picture where the father was missing. Therefore, the context information "the father is the agent of the 'missing' action" can be used to guide the parser to the appropriate analysis (Figure 4b).

8 Limitations

The main limitation of the approach is that it is only of benefit when strong context information is available (while writing and parsing). For example, if the topic of an essay is the only available context information, parsing would not profit from context integration. However, since the interpretation of learner utterances without context information is not reliable in general, Ott et al. (2012) recommend to collect language learner data with explicit task contexts.

Another disadvantage of the approach is its dependence on the verb of the sentence. If the link between the verb and the action in the A-Box cannot be established, the context information cannot be used to influence attachments to the verb. This could happen, e. g., if the verb is misspelled or inflected incorrectly or if it is mistaken for another verb. If the verb is missing completely, no verb-related error diagnoses such as case errors can be derived from the parser's output.

To make the matching of words to individuals more robust in general, a component needs to be added to the scoring mechanism for selecting referents in the A-Box (Section 6): This component could model typical errors such as false friends and dictionary errors and it could also deal with spelling errors more accurately, e. g., using an approach similar to King and Dickinson (2014).

If the content of the A-Box diverges from what the learner wants to express, integrating such context information either has no influence on syntactic parsing or, worse, the syntactic structure does not reflect the learner's intention. To mitigate these problems, several A-Boxes can be defined, and sentences will be parsed with each model separately. The model which yields the syntactic structure with the highest score can then be chosen as the most appropriate one.

9 Conclusions and Outlook

We have shown that the context of an utterance can influence the diagnosis of grammatical errors in several ways. Cases where the syntactic structure and the error diagnoses differ depending on the context have been identified and systematically categorized into confusion classes. We have proposed to use a model for integrating context information into syntactic parsing of language learner utterances to resolve these error-induced ambiguities.

In future work, we will evaluate this approach by parsing erroneous as well as well-formed sentences to find out whether context integration benefits parsing of faulty sentences and whether it deteriorates parsing of error-free sentences. For this purpose, we have collected texts written by learners of German, which describe the content of picture stories.

[5] TurboParser was trained on the first 100 000 sentences of part A of the Hamburg Dependency Treebank (Foth et al., 2014), the largest genuine dependency treebank for German.

References

Lene Antonsen, Saara Huhmarniemi, and Trond Trosterud. 2009. Constraint Grammar in Dialogue Systems. In Eckhard Bick, Kristin Hagen, Kaili Müürisep, and Trond Trosterud, editors, *Proceedings of the NODALIDA 2009 workshop Constraint Grammar and robust parsing*, volume 8 of *NEALT Proceedings Series*, pages 13–21.

Christopher Baumgärtner, Niels Beuck, and Wolfgang Menzel. 2012. An Architecture for Incremental Information Fusion of Cross-Modal representations. In *IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 498–503.

Sean Bechhofer, Frank van Harmelen, Jim Hendler, Ian Horrocks, Deborah L. McGuinness, Peter F. Patel-Schneider, and Lynn Andrea Stein, 2004. *OWL Web Ontology Language Reference*. World Wide Web Consortium (W3C). `http://www.w3.org/TR/2004/REC-owl-ref-20040210/`.

Emily M Bender, Dan Flickinger, Stephan Oepen, Annemarie Walsh, and Timothy Baldwin. 2004. Arboretum: Using a precision grammar for grammar checking in CALL. In *Proceedings of InSTIL/ICALL 2004 – NLP and Speech Technologies in Advanced Language Learning Systems*, Venice.

Adriane Amelia Boyd. 2012. *Detecting and Diagnosing Grammatical Errors for Beginning Learners of German: From Learner Corpus Annotation to Constraint Satisfaction Problems*. Ph.D. thesis, Ohio State University.

Thorsten Brants, Wojciech Skut, and Hans Uszkoreit. 1999. Syntactic Annotation of a German Newspaper Corpus. In *Proceedings of the ATALA Treebank Workshop*, pages 69–76, Paris, France.

Christian Fortmann and Martin Forst. 2004. An LFG Grammar Checker for CALL. In *Proceedings of InSTIL/ICALL 2004 – NLP and Speech Technologies in Advanced Language Learning Systems*, Venice.

Kilian Foth and Wolfgang Menzel. 2006. Hybrid Parsing: Using Probabilistic Models as Predictors for a Symbolic Parser. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL*, pages 321–328, Sydney.

Kilian Foth, Michael Daum, and Wolfgang Menzel. 2004. A broad-coverage parser for German based on defeasible constraints. In *KONVENS 2004, Beiträge zur 7. Konferenz zur Verarbeitung natürlicher Sprache*, pages 45–52, Wien.

Kilian Foth, Wolfgang Menzel, and Ingo Schröder. 2005. Robust Parsing with Weighted Constraints. *Natural Language Engineering*, 11(1):1–25.

Kilian A. Foth, Arne Köhn, Niels Beuck, and Wolfgang Menzel. 2014. Because Size Does Matter: The Hamburg Dependency Treebank. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Language Resources and Evaluation Conference 2014*, Reykjavik, Iceland. European Language Resources Association (ELRA).

Kilian A. Foth, 2006. *Eine umfassende Constraint-Dependenz-Grammatik des Deutschen*. Fachbereich Informatik, Universität Hamburg. URN: urn:nbn:de:gbv:18-228-7-2048.

Michael Hahn and Detmar Meurers. 2012. Evaluating the Meaning of Answers to Reading Comprehension Questions: A Semantics-Based Approach. In *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, pages 326–336, Montréal, Canada. Association for Computational Linguistics.

Birgit Hamp and Helmut Feldweg. 1997. GermaNet - a Lexical-Semantic Net for German. In *Proceedings of the ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, Madrid.

Trude Heift. 2003. Multiple learner errors and meaningful feedback: A challenge for ICALL systems. *CALICO journal*, 20(3):533–548.

Fred Karlsson, Atro Voutilainen, Juha Heikkilä, and Arto Anttila. 1995. *Constraint Grammar: A Language-Independent Framework for Parsing Unrestricted Text*. Mouton de Gruyter, Berlin and New York.

Levi King and Markus Dickinson. 2014. Leveraging Known Semantics for Spelling Correction. In *Proceedings of the third workshop on NLP for computer-assisted language learning at SLTC 2014, Uppsala University*, pages 43–58.

Julia Krivanek and Detmar Meurers. 2011. Comparing Rule-Based and Data-Driven Dependency Parsing of Learner Language. In *Proceedings of the International Conference on Dependency Linguistics (Depling 2011)*, pages 128–132.

Christine Köhn and Wolfgang Menzel. 2015. Towards parsing language learner utterances in context. Technical report, Fachbereich Informatik, Universität Hamburg. URN: urn:nbn:de:gbv:18-228-7-212.

André Martins, Noah Smith, and Eric Xing. 2009. Concise Integer Linear Programming Formulations for Dependency Parsing. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 342–350, Suntec, Singapore. Association for Computational Linguistics.

André Martins, Miguel Almeida, and Noah A. Smith. 2013. Turning on the Turbo: Fast Third-Order Non-Projective Turbo Parsers. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 617–622, Sofia, Bulgaria. Association for Computational Linguistics.

Patrick McCrae. 2009. A model for the cross-modal influence of visual context upon language processing. In *Proceedings of the International Conference Recent Advances in Natural Language Processing*, pages 230–235.

Patrick McCrae. 2010. *A Computational Model for the Influence of Cross-Modal Context upon Syntactic Parsing*. Ph.D. thesis, Fachbereich Informatik, Universität Hamburg.

Wolfgang Menzel and Ingo Schröder. 1999. Error Diagnosis for Language Learning Systems. *ReCALL*, (special edition, May 1999):20 – 30.

Joakim Nivre, Johan Hall, Jens Nilsson, Atanas Chanev, Gülsen Erygit, Sandra Kübler, Svetoslav Marinov, and Erwin Marsi. 2007. Maltparser: A language-independent system for data-driven dependency parsing. *Natural Language Engineering*, 13:95–135.

Erich Ohser. 1993. Der Schmöker. In *50 Streiche und Abenteuer*, volume 1. Südverlag GmbH, Konstanz.

Niels Ott and Ramon Ziai. 2010. Evaluating Dependency Parsing Performance on German Learner Language. In Markus Dickinson, Kaili Müürisep, and Marco Passarotti, editors, *Proceedings of the Ninth International Workshop on Treebanks and Linguistic Theories*, volume 9 of *NEALT Proceeding Series*, pages 175–186.

Niels Ott, Ramon Ziai, and Detmar Meurers. 2012. Creation and Analysis of a Reading Comprehension Exercise Corpus: Towards Evaluating Meaning in Context. In Thomas Schmidt and Kai Wörner, editors, *Multilingual Corpora and Multilingual Corpus Analysis*, Hamburg Studies in Multilingualism (HSM), pages 47–69. Benjamins, Amsterdam.

Veit Reuer. 2003. Error Recognition and Feedback with Lexical Functional Grammar. *CALICO Journal*, 20(3):497–512.

Margaret Rogers. 1984. On major types of written error in advanced students of German. *International Review of Applied Linguistics in Language Teaching*, 22(1):1–39.

Heike Telljohann, Erhard Hinrichs, and Sandra Kübler. 2004. The TüBa-D/Z treebank: Annotating German with a context-free backbone. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004)*.