MERLIN Annotation Structure Guidelines

VS 6

Please cite as: MERLIN project (2014): *Annotation guidelines*. www.merlin-platform.eu.

CONTENTS

0   INTRODUCTION

This document contains information for MERLIN staff concerned with manual annotations.

Chapter I gives an overview of the available text formats, of the tools to be used and the data management. Important extra information is given in the ANNEX to this document.

Chapter II describes the preparation of the MERLIN corpus and focuses on the tokenization procedures and piloting, thus excluding all working steps that came before, such as data collection, re-ratings, and transcriptions.

Chapter III sums up the reliability testing outline of the project.

Chapter IV gives an overview of the annotation structure, the annotation workflow, the role of documentation and communication which is central. It also gives a frame schedule with deadlines for the annotations.

Chapter V goes into the details of the single annotation layers in STAGE 1, in which minimal target hypotheses (TH1) and error annotations for grammar and orthography (EA1) are made for the whole corpus, and extended target hypotheses (TH2) and further annotations for a range of linguistic phenomena which are not necessarily errors (EA2) are annotated for a core corpus. The core corpus consists of texts that were rated A2 or B2 (see Chapter V.4 for the extra case of Italian).

Chapter VI mentions some aspects of interactions between single annotation layers. Most of them can also be found in Chapter V, however.

The **Annex** to these Guidelines is central for carrying out the work and knowing where to save and document what. It is also meant for short-term trouble-shooting.

I FORMATS, TOOLS, AND DATA

I.1 Available text formats

1) ORIGINAL TEXTS

The original learner texts are available in scanned pdf versions provided by the test institutions (UJOP and telc). They will not be published. They will be used for quality control of transcriptions. They shall be visible during annotation.

Available here: https://serifos.sfs.uni-tuebingen.de/svn-remote/merlin/corpus/scans/ (= Subrepository [svn] → corpus → scans)

2) TRANSCRIBED TEXTS

Texts were transcribed in xml mind. A number of transcription tags were used. These are documented in the transcription tagset and in the transcription guidelines. Transcriptions were carried through by telc and UJOP staff. Reliability was checked by UJOP and TUDD.

Tag	Description (also available in Italian and Czech)
<transcriber>	Name des Transkribenten
<author_id>	ID des Autors (= Testteilnehmers)
Tag	Kurz-Beschreibung
<anonymized>	Eine persönliche Information wurde ersetzt, also anonymisiert.
<hidden>	Eine persönliche Information wurde herausgenommen.
<ambiguous> <alternative>	Bei Unsicherheit des Transkribenten über eine Form (Buchstabe/Wort) im Text.
<citation>	Wenn der Autor aus Aufgabe/Inputtext abgeschrieben hat.
<closing>	Für Schlussformeln
<comment>	Kommentar des Transkribenten
<correction> <deletion> <insertion>	Korrektur des Textautors (!) im Text, mit 2 Spezifizierungs-möglichkeiten (Löschen/Hinzufügen)
<emoticon>	Emoticon-Verwendung
<emphasis>	Der Testteilnehmer hat etwas hervorgehoben.
<entity>	Eigennamen
<error> <original form> <target form>	Rechtschreibfehler im Text[1]
<foreign_word> attribute:foreign_language	Fremdsprachiges Wort (nicht: Fremdwort)

---

[1] The tagging of orthographic errors by transcribents can only be taken as an orientation for annotators; many classification errors were made as reliability checks of the transcriptions revealed.

<greeting>	Begrüßung
<image>	Bild
<list>	Auflistung
<originalForm>	siehe → <error>
<par>	Absatz
<symbol>	Symbol (etwa ein Pfeil)
<targetForm>	siehe → <error>
<unreadable>	Unleserlichkeit

Transcription-related documents
(available in Moodle: WP2: Data preparation → XML_mind transcription tool):

Transcription guidelines:

German/Italian: Tags_MERLIN-DE_FINAL-v2.doc
Czech: Transcription_guidelines_MERLIN-CZfin.doc
Anonymization: Anonymisierung_CZ.doc

3) ANNOTATED LEARNER PRODUCTIONS

Will be available in PAULA exchange format, Exmaralda format, MMAx format and Annis2.

Links available to the transcripts in
XML: https://serifos.sfs.uni-tuebingen.de/svn-remote/merlin/corpus/transcriptions/
PAULA format: https://serifos.sfs.uni-tuebingen.de/svn-remote/merlin/corpus/paula/
Exmaralda: https://serifos.sfs.uni-tuebingen.de/svn-remote/merlin/corpus/exmaralda/
Mmax: https://serifos.sfs.uni-tuebingen.de/svn-remote/merlin/corpus/annotation/mmax2/
One token per line: https://serifos.sfs.uni-tuebingen.de/svn-remote/merlin/corpus/one-tok-per-line/
Annis2: http://sifnos.sfs.uni-tuebingen.de/Annis-web/login.html

II.2 Tools & Data Management

Documentation of the annotation process

There are **two documents that accompany the annotation progress** and document any decisions taken:

References and guidelines	*Storage place*
FAQ-document	google docs
Annotation progress sheets – for all stages of	google docs
annotation	ITA
	GER
	CZE

FAQ document

This document is very important. The information contained here will be accessible by platform users, so please try to make it readable and write it in English. The Frequently asked Questions googledocs contains the RESULTS of discussions among annotators which are mainly carried out via email and Skype. It is important to keep track of all decisions that were taken and that could not be directly derived from the annotation scheme. It does not refer to single cases but to specific phenomena and how we went about annotating them, and why.

Annotation progress sheets for each language

These sheets are also very important. Here, the single files to be annotated are listed in an overview per step (e.g., Overview_TH1, Overview_EA1) and in single sheets for each annotator. This document is to be continuously updated so that we know where we are. Also, it serves as documentation; in case of critical single decisions, please document them here. Also, when annotating on level EA1, we document problems regarding TH1 that are detected. Please keep track of all changes applied to TH1 in the Annotation progress sheet.

TH1/TH2 annotation

Tools: Falko Excel Add-in, Exmaralda Excel Add-in

References and guidelines	*Storage place*
– **TH1-annotation-filemanagement.pdf**	moodle / WP5 / TH1 annotation tools, sources, guidelines
– **guidelines_Falko-Add-in.pdf**	moodle / WP5 / TH1 annotation tools, sources, guidelines
– **README_exmaralda_io_0.9.7.0.pdf**	moodle / WP5 / TH1 annotation tools, sources, guidelines

Error annotation (EA1 & EA2) and annotation of non-errors

Tool: MMax2

References and guidelines	*Storage place*
– **Technical manual for the MMax2 Tool**	moodle / WP5 / Error annotation tools and guidelines mmax2quickstart.pdf
– **MERLIN specific guidelines for MMax2**	moodle / WP5 / Error annotation tools and guidelines MMAX2 for MERLIN_guide.pdf
– **Guidelines to data-management with MMax2**	moodle / WP5 / Error annotation tools and guidelines MMAX-file-management.pdf
– **Guidelines for annotation of EA1 – TH2 – EA2**	Annex: "Guidelines for annotation of EA1 – TH2 – EA2"
– **tag span guide for EA1 (orthographic and grammar tags)**	moodle / WP5 / Error annotation tools and guidelines TagSpanGuideGRAM-ORTH_fin.xlsx
– **FAQ on the Mmax2 tool**	google docs – FAQ-document > sheet: FAQ MMax2

Detailed guidelines for error annotation including the reliability check procedure are given in Annex I.

II   PREPARING MERLIN ANNOTATIONS

II.1 Tokenization checks

After having completed transcriptions/digitisations, during this stage, we are mainly concerned with the preparation of a correctly tokenized version of all learner productions. In the annotation architecture, this will be represented in the level called "tokenized learner text" (see Table 1).

Tuebingen has already delivered automatically tokenized texts (available here: https://serifos.sfs.uni-tuebingen.de/svn-remote/merlin/corpus/one-tok-per-line/)**.**

The next step is the correction of these automatically produced versions which might contain errors. On this level, errors in tokenization are corrected **manually**. All tokenization errors (e.g.: superfluous space, wrong hyphenization) and all transcription errors must be deleted. Learner errors are not corrected in this stage, as the aim is to have a reliable representation of the learner's productions.

Guidelines are in Moodle (available here):
http://courses.sfs.uni-tuebingen.de/moodle/mod/folder/view.php?id=481**.**
Definitions of what we consider a token are to be found here:
https://docs.google.com/spreadsheet/ccc?key=0AoiqtFyBRJytdFpLMDE1RkNKcF9XQ0otOEU3em9VaWc#gid=0

(Deadline: January 20th 2013)

However, for "emergency cases" of further errors that remain undetected before the annotation process starts, we have the ctok level in which we can correct these errors (see III.6 and Table 1).

II.2   Piloting

A piloting must be carried through before the annotation process starts. **All levels of annotations must be tried out for all languages,** and all annotators must be acquainted with the complete annotation scheme. A particular concern is the feasibility of writing target hypotheses.

Should there be very problematic issues in annotation, these will be discussed, and the annotation scheme will be modified if needed.

During the piloting stage, all annotators use the preliminary annotation structure in excel format. As draft serves an excel table with several worksheets which contain a short guideline for the piloting, general information about the selected texts (one for each level in each language = 12 texts), working plan and the texts themselves.

(available in Moodle: http://courses.sfs.uni-tuebingen.de/moodle/mod/folder/view.php?id=482).

Comments Piloting Procedures for TH2

Special measures of quality control are taken for TH2 which will be piloted in a two-step procedure. In a pre-piloting, the team tries out the TH2 rules and gives feedback to the co-ordinator who then adapts the rules (this document). 5 productions per language are pre-piloted.

For the actual piloting, 6 texts per language are annotated in Excel and discussed by the language team. If necessary, guidelines will be adapted again.

Deadline: End of July 2014

Comments Piloting Procedures for EA2

The texts chosen for TH2 pre-piloting also receive the EA2 annotation tags in order to check the consistency and transparency of the annotation scheme which will be modified if necessary. Deadline: End of July 2014.

The second piloting round takes place toward the end of the TH2 annotations (early autumn 2014).

Summing up:

Before annotation starts

1) Check tokenization

2) complete Annotation Scheme(s) !

3) add language-related information to this document

4) Piloting of AS in WP5, including target hypotheses

5) Revision of AS by WP5

6) guidelines for annotation (see Annex I "Guidelines for annotation of EA1 – TH2 – EA2")

7) Tryout of annotation tool

III   INTER-CODER RELIABILITY IN MERLIN

Reliability checks are carried out for each level and language. Annotators discuss reliability files which are always annotated independently first (in Excel), and they agree upon a final annotation version which is committed to the "gold version" folder of their language. Further reliability checks are carried out without the knowledge of the annotators. It is important to do the checks during annotation, not afterwards! Difficult aspects will thus be detected and can be corrected at an early stage of annotation.

All in all, 5% of productions per level and language are annotated by all annotators in STAGE 1 and also in STAGE 2. Inter-coder agreement will be calculated on basis of a) appearance of text (dichotomous) and b) agreement of exact matches in choice and span of tag.

Where possible, the files that undergo reliability checks in stage 2 correspond to the ones in stage 1. Text 1 files chosen for reliability annotation, however, were based on test levels, whereas the stage 2 sub-corpus is based on fair averages which are more meaningful for users. Thus, not all files that go into stage 2 reliability checks are identical to the ones in stage 1. The annotations of vocabulary, sociolinguistic appropriateness, and pragmatics is not as straightforward as EA1 annotations, thus it might prove more problematic to arrive at unambiguous annotations.

IV   ANNOTATING THE MERLIN CORPUS: OVERVIEW

IV.1 MERLIN annotation structure

Graphic 1 gives an overview of the structure of annotation in Merlin. Table 1 shows all levels of annotation planned to be integrated in MERLIN. Details can be found in the following paragraphs.
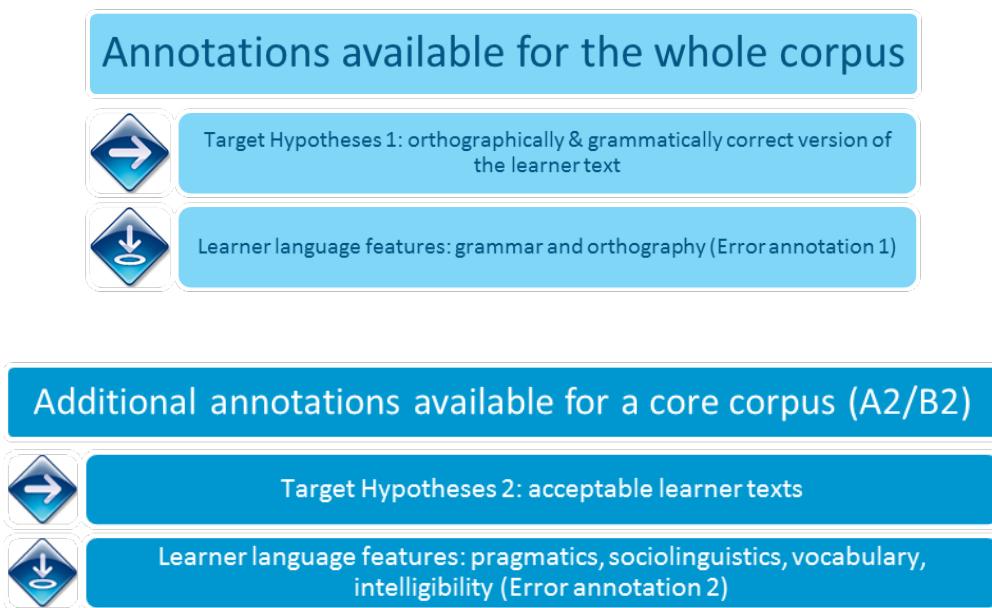
Graphic : MERLIN annotation architecture

STAGE 1: Annotation of Target Hypothesis 1 (TH1) and Error Annotation 1 (EA1). We will proceed level by level. Differing from the proposal, we *start with the lower competence* levels. Reason: the higher number of learner errors at these levels enables us to detect problems with the annotation tools as early as possible.

STAGE 2: Annotation of TH2 and EA2. First, TH2 will be written for selected texts rated levels A2 and B2, based on the fair average ratings (not on the original test levels). The same sub-corpus will receive EA2 annotations.

In EA2, all phenomena that were not part of EA1 will be annotated, including: intelligibility, vocabulary, sociolinguistic appropriateness, pragmatics, coherence/cohesion, regardless of whether they refer to errors in learner language or to other aspects.

Resulting annotation tiers in MERLIN

Tok				
Ctok				
ZH1² incl. levels ZH1Diff and ZH1spec				
EA1_lev1				
EA1_lev2				
EA1_lev3				
EA1_tlm				
ZH2 incl. level ZH2Diff and ZH2spec				
EA2_lev1				
EA2_lev2				
EA2_lev3				
EA2_tlm				
Comment tier				

Table 1: structure of annotations in Merlin

Abbreviations:

| Tok	=	automatically tokenized and manually checked learner text
Correctly represented learner production (see III.6)		
Ctok	=	level for emergency corrections of tokenization/transcription, see III.4!
Stage I (orthography & grammar, paragraphs III.2.1 - III.2.3)		
TH1(ZH1)		Target Hypothesis 1 (complete, corrected learner text)
ZH1Diff	=	level for marking differences between TH1 and ctok -> Falko
ZH1spec	=	level for marking speculative hypotheses (1=uncertain, but inferable; 2=TH1 is a wild guess, 3= no TH1 possible as partly not comprehensible)
EA1_lev1,2,3		Error annotation 1, specified on level 1-3 (see Annotation Scheme)
EA1_tlm		target language modification (see Annotation Scheme)
Stage II (vocabulary, sociolinguistic, pragmatic errors, paragraphs III.2.4 & III.2.5)		
TH2(ZH2)		Target Hypothesis 2
EA2_lev1,2,3		Error annotation 2, specified on level 1-3 (see Annotation Scheme)
EA2_tlm		target language modification (see Annotation Scheme)

---

[2] ZH is the naming convention for target hypotheses in the Falko-Excel macro. As the MERLIN-team decided to use this tool for the annotation of the target hypotheses, in our data TH1, TH2 and the related tiers for marking differences carry the name "ZH".

IV.2 MERLIN annotation workflow

Graphic 2 gives an overview of the steps to take in order to arrive at the MERLIN annotations:

Annotation process in Merlin

PREPARATION OF ANNOTATION
(automatic tokenization, manual check of tokenization and transcription)

PILOTING OF ANNOTATION SCHEME (AS) & TARGET HYPOTHESES (TH) GUIDELINES

Stage I: orthography and grammar

Target Hypothesis I: Rewrite complete productions without grammatical and orthographic errors (→ Falko)
Error Annotation I: Annotate grammatical and orthographic errors according to AS

error-related

AUTOMATIC ANALYSES (POS; parser…)

GOLDEN VERSIONS (TH1/EA1)

Stage II: vocabulary, sociolinguistic, coherence/cohesion/pragmatics/intelligibility

Target Hypothesis II: Rewrite aspects of productions; aim: acceptability, levels A2 & B2 (→ Falko)
Error Annotation II: Annotate sociol., lexical, pragmatic, coherence/cohesion phenomena and intelligibility according to AS, levels A2 & B2

error-related and not error-related

Graphic 2: Stages in the annotation of the MERLIN corpus

WP5 & WP6 co-operation

The annotations require a sophisticated working procedure, especially because of the quality controls *inside* each annotation step and the conversions necessary *between* annotation levels. WP5 people co-operate very tightly with the computational linguists from WP6 as the realization of the platform is strictly dependent on the timely fulfilment of work duties from both manual and automatized annotations. For example, in order to step **from STAGE 1 to STAGE 2**, we need to …:

➢ create gold versions of EA1 annotations in MMAX

➢ convert texts with TH1 and EA1 from MMAX2 to PAULA

➢ create new TH2 EXB files to use with the FALKO add-in

➢ annotate TH2

➢ convert FALKO XML files to PAULA

- ➢ convert the PAULA files to MMAX2

- ➢ update MMAX2 styles

- ➢ create MMAX2 annotator directories for EA2

- ➢ annotate EA2

An overview is given by Graphic 3 and Graphic 4:
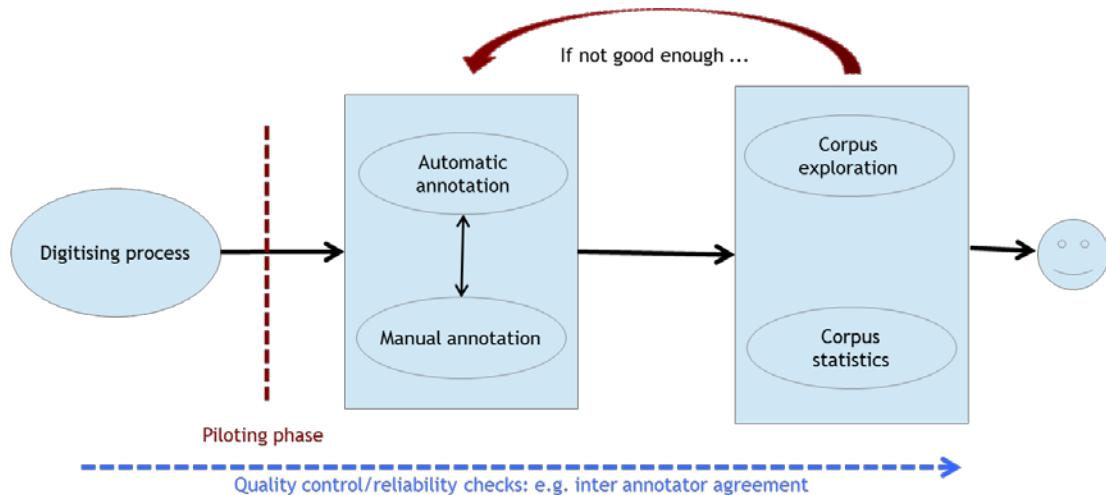
Graphic 3: Workflow in MERLIN, WP5 perspective (graphic by Andrea Abel)
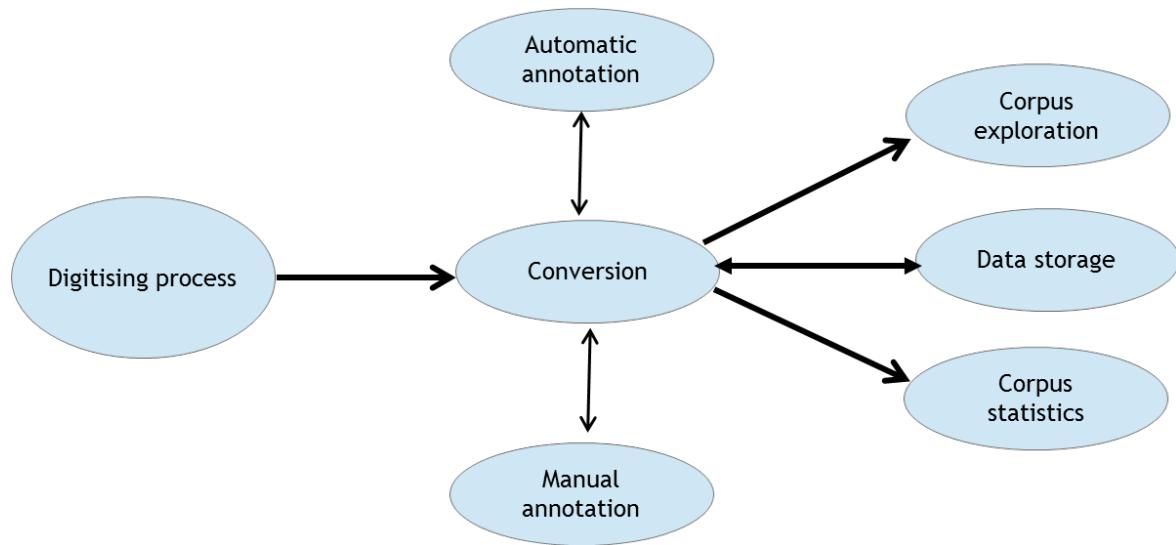
Graphic 3: Workflow in MERLIN: computational linguist view (graphic by Andrea Abel)

IV.3 The role of documentation and communication

To guarantee high quality of annotations, documentation plays a vital role.
All questionable/difficult decisions are documented (see FAQ in Google docs:
https://docs.google.com/spreadsheet/ccc?key=0AouRwmLOZeQudEw2ZmJ0dGtzQm5Nc0c1U3NjTVdm
OHc#gid=1). The documentation must be very well structured so that the others can make use of it as a
reference document. It will be **publicly accessible** and part of the MERLIN manual for users.

The **annotation progress** is also continuously documented, in the Annotation Progress Sheets for each
language (e.g. AnnProgGER). Like this, everybody knows how far annotation has gotten, and
responsibilities for the annotations are clear.

More information on where to document what is given in the Annex to this document.

Communication: Annotators annotating the same language communicate very tightly about their
progress & possible difficulties (email & skype). To ensure for comparability across languages, we
organize regular skype meetings among all annotators. If any annotation issues regard all languages,
common discussion among WP5 are incentivated. If there are technical problems, WP6 is integrated and
all WP5 members are included in the communication.

IV.4 SCHEDULE

The MERLIN annotation architecture is complex. Therefore, we decided to adjust the deadlines
according to the progress and the general conditions (staff; tools) at a given point of time. The deadlines
for single (micro-) aims are defined throughout the annotation project and communicated via email.

However, there are some framing conditions that force us to keep the following deadlines:

STAGE 1-related deadlines:
1) TH1 corrected and revised for core corpus - **July 23, 2014**
2) TH1 corrected and revised for whole MERLIN corpus – **September 15, 2014**
3) EA1 in gold folders for core corpus – **July 23, 2014**
4) EA1 in gold folders for whole MERLIN corpus – **September 15, 2014**

STAGE 2-related deadlines:
1) TH2 for core corpus ready – **September 30, 2014** (for German: September 30, 2014, as there are
twice as many files)
2) EA2 for core corpus ready – **November 20, 2014** (for German: November 30)

V   Annotation in MERLIN: details of single annotation levels

V.1   Target Hypotheses – General Rules and Information

For the description and classification of learner language errors, it is necessary to define at least one target hypothesis, i.e. an explicit assumption about what the annotator(s) assume(s) the learner intended to say. This process is subjective, but inevitable. In order to improve the annotation of target hypotheses (TH) and also to enhance transparency and comparability, most of the definitions and procedures of the Falko corpus are adopted by the MERLIN team.

For reasons of practicality and reliability, the **two-step structure** that Falko uses will be adopted by MERLIN. That means that after having prepared the texts for annotation, first a **target hypothesis for orthographic and grammatical errors (TH1)** will be written, and these kinds of errors will be annotated (EA1) [3]. Then, an ext**ended target hypothesis (TH2) that refers to aspects of sociolinguistic, lexical, and pragmatic deviation**s from what would be expected from a native speaker will be written. EA2 consists of tags from these fields.

Additionally, using the surface error markers as used in the Falko-corpus (e.g. for deletions, movement, insertions, changes) the two tiers **ZH1Diff and ZH2Diff** are created in order to illustrate all differences between the learner's text (ctok) and TH1/TH2, i.e. the changes that have been made by the annotator.

The TH are not part of the error annotation, but a prerequisite for it: „eine implizite Annotation von Abweichungen der Lernertexte von einer postulierten Standardvariante, vor deren Hintergrund eine Fehlerannotation möglich wird und die auf dieser aufbauen kann. Für eine kompakte Darstellung auf Englisch siehe Reznicek et al. (im Druck)." (Falko Handbook et al. 39). [engl.: "an implicit annotation of learner deviations from a postulated standard language variety constitutes a basis and makes the annotation of errors possible in the first place (…)]

The error phenomena that will be tagged in MERLIN are defined and described in the annotation scheme(s) (AS). The technical proceeding is described in "guidelines_Falko_Add-in" (see II.4).

General rules for writing Target Hypotheses

These rules are valid for TH1 and TH2. They are to be followed at all times.

The following examples illustrate the basic rules i.e., guidelines for the creation of target hypotheses.

---

[3] In a small number of cases grammatical errors will be corrected only on TH2 level as a consequence of following the FALKO guidelines for TH1-annotation. This is the case for tense and mood conflicts as well as the wrong usage of definite/indefinite articles that can be observed in context only. Nonetheless those errors are to be annotated with the appropriate grammar tags.

Please note that all examples are taken from Falko. *ZH* stands for target hypothesis (TH) and defines which change of the learner text is considered appropriate (green color). The line with red coloring on the other hand shows an example of TH1 creation that should be avoided.

ZH1/2Diff describes differences between TH1 (TH2) and the ctok tier using the Falko surface error markers (see III.4)

1   Minimal Invasiveness (what needs not to be changed will not be changed)

ctok	Frauen	konnten		solchen		geselschaftlichen	Zustand	verändern
ZH1	Frauen	konnten	einen	solchen		gesellschaftlichen	Zustand	verändern
¬ZH1	Frauen	konnten		solch	einen	gesellschaftlichen	Zustand	verändern

(fk024_2006_07)

2   Move instead of Change (constituents will not be exchanged, constituents are always moved to empty token positions in the TH)

ctok	Man	hat		ihr	es	geglaubt	.
ZH1	Man	hat	es	ihr		geglaubt	.
¬ZH1	Man	hat		es	ihr	geglaubt	.

3   For German: Move left rather than right

ctok	Man	hat		ihr	es		geglaubt	.
ZH1	Man	hat	es	ihr			geglaubt	.
¬ZH1	Man	hat			es	ihr	geglaubt	.

4   Small Movements: If a token has to be moved to make a TH1, keep the move as short as possible.

ctok		Die	Frauen	haben		die	Macht	auch	.
ZH1		Die	Frauen	haben	auch	die	Macht		.
¬ZH1	Auch	die	Frauen	haben		die	Macht		.

5   Move as little constituents and tokens as possible

ctok	Es	hat					Geld	seiner	Firma	in	Polen		gebracht	.
ZH2	Es	hat						seiner	Firma	in	Polen	Geld	gebracht	.
¬ZH2	Es	hat	seiner	Firma	in	Polen	Geld						gebracht	.

6   Merge and split tokens: If tokens are merged/split, the cells will be merged/split, too.

16

ctok	zum		Frau
ZH1	zu	der	Frau

7   If a learner writes something twice, it is deleted from the right.

ctok	es	nicht	es	nicht
ZH1	es	nicht		
¬ZH1			es	nicht

8   If the construction of the TH involves merging/splitting the original version, if possible these should be annotated as overlaps.

ctok	ob	Forschung	frei	oder	mit	einem	gegebenen	Ziel	sein	solle	.
ZH2	ob	Forschung	frei	oder	zielgerichtet				sein	sollte	.
ZH2Diff					MERGE					CHA	

V.2 Stage 1, TH1: MINIMAL TARGET HYPOTHESIS

Aim: create grammatical & orthographical correctness: This annotation level creates a version of the text, which is usable for automatic processing (Luedeling et al. 200x: 38). The aim of TH1 is to create a parsable structure and a version of the text that is in line with the rules of Italian, Czech, and German orthography and morphosyntax. Lexical errors are not corrected here, but in TH2.

The TH1 contains the complete learner production. If errors appear, minimal changes to the learner texts are carried through (follow general rules III.2.1 and specific rules, see below). These changes regard grammar and orthography only.[4] A target hypothesis is to be created if an orthographic, syntactical, or morphological error appears that makes the learner's utterance incorrect. If there is no error, the TH1 level contains a repetition of the learner text in the original version, including all lexical, sociolinguistic, pragmatic errors (these will be taken care of in stage 2, see below).

The aim of TH1 is to have a text version that is correct on the grammatical & orthographical level. Thus, the cases in which we need a TH1 are dependent on the German/Italian/Czech grammar and orthography rules. It is possible that there will be target hypotheses for cases in which grammar and orthography rules might have been hurt for which we decided not to make an error annotation. As the MERLIN AS scheme is very comprehensive, this is not to be expected to happen very often[5].

Please note that every sentence should be considered on its own, i.e., in isolation, during grammar and orthography annotation (TH1), without taking into consideration the context it is written in. In this way, consistency between annotators can be guaranteed and the focus on pure grammar and orthography is ensured.
As a consequence of this approach in a limited number of cases, in which there is a grammatical correct sentence and it appears only from the context that the learner has made a grammatical error, this grammatical error will be corrected only in TH2, but annotated with EA1 tags[6]. This is the case for some incidences of tense and mood conflicts, wrong choice of definite/indefinite articles and the wrong choice of a preposition to be observed in context only. To ensure consistency between anticipated errors and target hypotheses in conflict cases a TH2 is written for those cases immediately after TH1 (see guidelines_Falko_Add-in).

The MERLIN team decided to mark speculative hypotheses in a special tier "ZH1spec" to be able to detect cases where deviating hypotheses are most probable, e.g. incomprehensible or very uncertain learner sentences and missing words.

Why do we need the TH specifications in addition to the EA? The AS as well as the grammar and orthography rules of each language dictate **IF** to write a TH. However, it is good to have general and

---

[4] Like in the Falko corpus, language that is not the L2 will be translated in the TH1 level to ensure that the language material used and its morphosyntactic properties are according to the target language. In MERLIN, additionally a transcription tag was used for this phenomenon.

[5] Those cases are documented in the FAQ-document/sheet: "changes on TH1 without tag in EA1"

[6] for procedure see: Annex I

specific rules about **HOW** to do that, particularly if we take into account the subjectivity of that process. Furthermore, this level is needed for automatic analyses.

Some rules for writing TH are given below. Again, the majority of the examples stems from the Falko corpus; they present examples of how annotation should proceed (green color) and what should be avoided (red color).

First, general, cross-language rules are cited. Then, rules for German, Czech and Italian are described.

Rules for writing TH1, all languages:

Orthography

1   If whole words are written in capital letters in order to emphasize something, these will be rewritten in standard capitalization. No error annotation. See Falko handbook p. 43.

If the **whole learner text** is written in capital letters, we rewrite it in TH1 (and TH2) following the orthographic rules of the concerned target language, i.e. we write *Haus* instead of *HAUS*, *movimento* instead of *MOVIMENTO* etc. The change is not commented, i.e. we do not annotate the change on EA1 level. The rule applies only for the capital letters, not for other orthographic errors.

Grammar

1   Exception to the rule "minimal movement 1": If the smallest move possible is very different from the learner's original text and at the same time not very representative of the target structure, it is possible to hurt the 'minimal movement' rule (see above, see Falko handbook p. 43).

Abweichung von minimaler Bewegung
Führt eine minimale Bewegung zu einer Struktur, die einerseits sehr stark von der Originalsyntax abweicht, die intendierte Zielstruktur aber andererseits nicht gut abbildet, so darf die minimale Bewegung verletzt werden.

ctok		In	Neu	Seeland	auch	haben		wir	eine	Prime	Minister
ZH1	Auch	in	Neuseeland			haben		wir	eine	Preministerin	
¬ZH1		In	Neuseeland			haben	auch	wir	eine	Preministerin	

2   Exception to the rule 'minimal movement' 2: If there is a coordinate structure, and the minimal movement would separate the coordinate structure, a longer movement is preferred.

ctok	obwohl	ich	erkläre	meine	Meinung		und		von	meinen	Freunden	
ZH1	obwohl	ich		meine	Meinung		und	die	von	meinen	Freunden	erkläre
¬ZH1	obwohl	ich		meine	Meinung	erkläre	und	die	von	meinen	Freunden	

3   Verbal nucleus: If the verb does not correspond to its arguments, keep the verb, and adapt the arguments. If obligatory objects are missing, dummies are inserted. Further examples Falko handbook p. 45.

ctok	dass	unsere	Zivilisation	sich	aus	Gruppen	besteht	.
ZH1	dass	unsere	Zivilisation		aus	Gruppen	besteht	.
¬ZH1	dass	unsere	Zivilisation	sich	aus	Gruppen	zusammensetzt	.

4   Hierarchy of nominal agreement: In case of lack of agreement between determiner, attribute and nominal nucleus, focus is on the nucleus. After that, decisions are taken in favour of the determiner. Highest priority = syntactic nucleus (lexical head), followed by determiner, followed by attributes.

ctok	keine	Arbeit	bekommen	,	die	ihrer	Qualifikationen	entspricht	.
ZH1	keine	Arbeit	bekommen	,	die	ihren	Qualifikationen	entspricht	.
¬ZH1	keine	Arbeit	bekommen	,	die	ihrer	Qualifikation	entspricht	.

5   In case of missing articles, insert the article that best fits the context (be it a definite or an indefinite article. This example is taken form the MERLIN corpus since the corresponding Falko rule (Version 2.01) is not followed here.

ctok	Meiner	Meinung	nach	Kann	man		solche	Regelung	nicht	akzeptieren	.
ZH1	Meiner	Meinung	nach	Kann	man	eine	solche	Regelung	nicht	akzeptieren	.

6   Conflicts of tempus and mode consistency: if the semantic or pragmatic structure of an utterance requires a certain tempus or mode, and that is not to be observed, no correction is made in TH1. Exception: If the conflict is observed within a compound sentence, a TH1 is created. In this regard, MERLIN deviates from the corresponding rule in the Falko handbook (version 2.01).

ctok	Und	wenn	mich	jemand	darüber	fragt	,	erzälte	ich	Gerne
ZH1	Und	wenn	mich	jemand	dazu	fragt	,	erzähle	ich	Gerne

7   Dates: Some tasks require the learners to write an informal/formal letter where they also have to include a date. The position of the date in a learner's text will neither be changed nor annotated. The wrong format of dates is not expected to be corrected nor annotated.

8   Errors in proper names (including addresses): a correction of proper names in TH1 is expected only for toponyms and names of persons / institutions, if the referred entity is clear. For manual error annotation only a) names of countries, continents, oceans and b) names of regions / federal states (endonyms) of the related target language country are considered[7].

Special cases referring to different levels of language description

9   Missing words: Missing content words will be added in cases where they are necessary to ensure a syntactically well-formed sentence (elliptic sentences are excluded). The insertion is marked in the "ZH1spec" tier as 1 (uncertain but inferable from the context) or 2 (insertion is a wild guess)[8]. The example below is taken from MERLIN and illustrates an instance where the missing word can be inferred from context.

Ctok	…	wie	gehst	:	sind	beide		?
ZH1	…	wie	gehts	:	sind	beide	gesund	?
ZH1spec							1	

10 Incomprehensible (parts of) sentences: For apparently incomprehensible parts of sentences annotators should try to find an interpretation and mark the speculation in the "ZH1spec". If this is not possible, for the word, clause or sentence in question an <-incomp-> is inserted into TH1 and the span is marked in "ZH1spec" with a <3> (for incomprehensible).

11 ambiguous readings: If there are two possible readings of a word the most probable alternative is inserted into ctok and thus copied to "ZH1". The other alternative is to be inserted into the existing transcription tier "tok_ambiguous_alternative1". For the procedure and an example see "guidelines_Falko_Add-in".

12 unreadable parts of sentences: Unreadable (parts of) sentences have been tagged in the transcription as <-unreadable->. This tag is preserved also in TH1.

13 words from other languages: words from other languages than the L2 will be translated if possible (exception: the word/phrase in question is used for explanation/specification of the L2 equivalent and therefore written in brackets or quotation marks. (E.g. *Due anni fà mi sono diplomata nel **"gymnasium"** che corrisponde al **liceo**.*). The error is to be annotated in the EA2 stage only.

Ctok	…	Das	new	Baby	

---

[7] for Czech also names of Czech towns (endonyms) and Czech names of foreign towns / regions (exonyms), eg. Londýn for London, Drážďany for Dresden are considered.
[8] The marking in „ZH1spec" is not mandatory, if the insertion was not speculative at all.

ZH1	…	Das	neue	Baby

14 wrong choice of connectors: please differentiate: if the choice of the connector leads to a syntactically wrong sentence, it is to be corrected on TH1 but a semantically/pragmatically wrong choice on TH2 only (GER: LV=TH1: "*Ich weiß nicht, dass Du kommst.*" → TH2 "*Ich weiß nicht, ob …*").

15 errors in superlative/comparative forms: grammatically erroneous forms are to be corrected on TH1. Inappropriate usage of correct comparative /superlative forms where no syntactical conflict occurs is to be corrected on TH2 only.

16 word formation: words that are morphosyntactically correct (eg. *Bewerbergespräch*, *corraggiare*), i.e. the single morphemes (including derivational suffixes) are correct even if the combination itself does not exist, are not corrected on TH1 (in TH1 only correction of orthographical errors); word formation errors including missing or wrong "Fugenelemente" are to be corrected in TH2.

ctok	Bewerber	Gespräch
ZH1	Bewerbergespräch	
ZH2	Bewerbungsgespräch	

ctok	Arbeitplaz	němĕčtínu
ZH1	Arbeitplatz	nĕmĕčtinu
ZH2	Arbeitsplatz	nĕmčinu

Rules for writing TH1 for German only

Orthography

1. The spelling is to be corrected according to the official German spelling rules valid from August 2006.

2. Swiss spelling is accepted. ("Ein grosses Haus; <ss> instead of <ß>"). Note for error annotation: If the learner uses the grapheme <ss> consistently in the whole text, it is not counted as an error. However, if the learner uses <ss> and <ß> inconsistently (exp. Straße and Strasse in one text), the Swiss spelling counts as an error.

3. If "Du/du" and "Ihr/ihr" (informal address) is used for addressing someone in a letter both the spelling with capital letter and with small letter is accepted and will not be corrected in TH1. This applies also if the learner uses <Du> and <du> inconsistently.

4. If Umlaute are written ue, ae, oe, they are rewritten as Umlaut. No error annotation.

5. Merging of preposition and article that are used in spoken language will not be corrected in a TH (exp. *sich durchs Leben schlagen*, Falko, p. 43).

6. Handling of *geht's* in ctok and TH1: *geht's* will <u>not</u> be changed in TH1 to *geht es*. In case *geht's* represents three tokens (i.e., three separate lines) in ctok, it should be merged to one token both on ctok and TH1 level (see below).

Tok	Wie	Geht	'	S	?
Ctok	Wie		geht's		?
ZH1	Wie		geht's		?

Grammar

7. If there is a German subordinate clause introduced by a connector, followed by a wrong word order, the word order must be corrected in TH1. This is to be done even if it would cause less change to replace the connector chosen by the learner with another one. See next example (and Falko handbook p. 43).

8. Word order in subordinate clauses: V2-Stellung is corrected in TH1. Falko handbook, p. 44. The example below is taken from MERLIN.

ctok	Wenn	hast	du	Lust			nach	Dortmund		fahren
ZH1	Wenn		du	Lust	Hast	,	nach	Dortmund	zu	

9. Exception to the rule 'move left': Keep 'linke Satzklammer'. In main clauses, the finite verb is treated as linke Satzklammer which is not moved. This rule has higher priority than the 'move left' rule.

ctok	Heute		Männer	gehen		nicht	oft	einkaufen	.
ZH1	Heute			gehen	Männer	nicht	oft	einkaufen	.
¬ZH1	Heute	gehen	Männer			nicht	oft	einkaufen	.

10. Rechte Satzklammer: Finite verb may be moved. Falko handbook p. 44.

ctok	was	ist	darüber	gemeint		.
ZH1	was		damit	gemeint	ist	.
ZH1Diff		MOVS	CHA		MOVT	

11. If it is impossible to add arguments to the verb, a prefix will be inserted or a particle will be added.

Änderungen am Verb
Können dem Verb keine Argumente hinzugefügt werden, wird das Verb präfigiert oder mit Partikel versehen.

> ZH1: Man sagt, dass die Arbeitsbedingungen für Arbeitseinsteiger nicht verbessert wer-
> den können, es sei denn, der Staat **passt** neue arbeitsmarktregelnde Gesetze **an**.
> (fk016_2006_08)

> ctok: Man sagt, dass die Arbetsbedingungen fuer Arbeitseinsteiger konnen nicht verbes-
> sert werden, es sei denn der Staat <u>passt neue Arbeitsmarkt regelnde Gesetze</u>.

12. For the government of case by prepositions, Grammis (Breindl 2000; http://hypermedia.ids-mannheim.de/) is followed. Tendencies to use the dative case are accepted and will not be changed in the TH1.

> **Kasusrektion durch Präpositionen**
> Kasusrektion durch Präpositionen wird gemäß der Darstellung in Grammis (Breindl 2000) be-
> handelt, hierbei werden auchprogressive Entwicklungen hin zum Dativ akzeptiert

> ctok: Während den sechziger Jahren, gab es eine grosse Debatte über Porn.
> ZH1: Während den Sechzigerjahren gab es eine große Debatte über einen Porno.
> (BNG2-2010-11-120)

13. If it is still grammatically acceptable, do not add or delete correlates. In the first example (from Falko) given below, the correlate *es* is redundant and therefore has to be removed. The second example (from MERLIN) shows a use of a correlate which is acceptable.

> **Korrelate**
> Es wird vermieden, Korrelate hinzuzufügen und zu entfernen, wenn dies noch grammatisch
> möglich ist.

> ctok: Wenn man die Argumente abwägt, wird es klar, dass Kriminalität oft sich nicht
> auszahlt.
> ZH1: Wenn man die Argumente abwägt, wird klar, dass Kriminalität sich oft nicht aus-
> zahlt.

ctok	Ich	habe	deinen	Brief	gelesen	.	Es	freut	mich	sehr	.

14. Word order in the Mittelfeld: To be corrected only if ungrammatical structures are observed.

ctok	Er	hat		meiner	Firma	es	gesagt	.
ZH1	Er	hat	es	meiner	Firma		gesagt	.
¬ZH1	Er	hat		meiner	Firma	es	gesagt	.

15. Infinitive construction with *zu*. If there is a sentence with the missing or redundant complement *zu*, it will be corrected, i.e. added/omitted in TH1 since it constitutes an error in the use of grammatical complements. (Example below is taken from MERLIN)

ctok	Dies	Zwei	Frage	richtig		beantworten	ist	nicht	einfach	.
ZH1	Diese	Zwei	Fragen	richtig	Zu	beantworten	ist	nicht	einfach	.

16. Word order in the "Nachfeld": Worstellung im Nachfeld:
a) The word order shall be corrected (and annotated as an error) if the following syntactical items appear in the „Nachfeld": subject, object in the dative/accus./genitive case; nouns that are part of „Funktionsverbgefüge"; predicatives; "Pronominaladverbien", and adverbial complements that do not form a prepositional phrase
b) exception → not to be corrected:
adverbial complements and adjuncts in the "Nachfeld"-position that are a prepositional phrase

ctok	Ich	Bin		glücklich	jetzt	.
ZH1	Ich	Bin	jetzt	glücklich		.

There are some details in the Falko handbook which are not mentioned here because they are considered to be of less importance for us. Anyway, should these phenomena appear during TH1, we will adhere to the Falko rules (Handbook, pp 43-50).[9]

Rules for writing TH1 for Italian only

Grammar

1. In cases where the subject is not explicit in the sentence but inferable from the context, if the verb form is grammatically correct, the context is not considered and the verb is not corrected in TH1 as G_Agr. Annotation in EA2: as C_coh_ref. Examples: "*Secondo che tipo lavoro voule fare?*", "*Come stai? Che cosa fai? Ha anche molto lavoro?*" (In both cases the subject inferred by the context is "tu" and the speaker was addressing to him in an informal way).
2. inconsistent capitalization of pronouns for addressing someone in a letter: the capitalization is corrected according to the form that was mainly used in the learner text. If capitalization is not obligatory, no error annotation will take place.

---

[9] Konkurrierendes 'dass' ; linke Satzklammer erzeugen.

Further Italian specific rules that are related to TH1 are specified in the Annotation scheme (AS) in: G_Inflect_noun_inexist (for cases like: "un'aiuta" instead of "un aiuto", "passeggiati" instead of "passeggiate") and G_prep (articulated vs. simple prepositions).

Rules for writing TH1 for Czech only

Grammar

1. In cases where the subject is not explicit in the sentence but inferable from the context, the Czech team consequently treats those cases as agreement errors (and thus corrects them on TH1), because those errors aren't perceived to be a reference error. Nevertheless, the reference error is also to be annotated in EA2**.**
2. Informational structure of a sentence/inappropriate topic-focus articulation: A word order that is grammatically acceptable but inappropriate as for the informational structure is not corrected on TH1.

3. Usage of upper and lower index quotation marks: According to the decision taken for the CZesL corpus the wrong format of quotation marks isn't considered neither for TH1 correction nor for error annotation.
4. Inconsistent capitalization in addresses: If both "Tvůj" and "tvůj" is used for addressing someone in a letter, the capitalization is corrected according to the form that was mainly used in the learner text. If capitalization was obligatory, missing capital letters are to be tagged as errors.

For further questions, a Czech version of the FAQ-document is available. (Veronika will summarize major items in English and add them).

V.23 STAGE 1: ANNOTATING ORTHOGRAPHY AND GRAMMAR (EA1)

While TH1 delivers a descriptive account of what the learner text might have looked like if no orthographic and grammatical errors had been made, the error annotation (EA1) tries to classify these errors concretely. Orthographic and grammatical errors are tagged as described in the annotation scheme. One TH can contain several error tags.

Here, only a short overview of errors to be tagged will be given. A more comprehensive list with a decision rationale and examples can be found in the annotation scheme (AS). For more detailed guidelines please see Annex I (Guidelines for annotation of EA1 – TH2 – EA2).

In MERLIN the complete error annotation structure including the EA levels 1&2, 3 if applicable, and the EA_tlm are applied.

The EA (EA1 and EA2) are structured as follows:

a) **General linguistic subfield** (level 1) (GRAMMAR/ORTHOGRAPHY)[10]

b) **error classification I** (level 2) (e.g. agreement error, word order error, punctuation error, … For a complete list see AS)

c) **error classification II** (level 3) (e.g. missing agreement between subject and predicate. For a complete list see AS)

d) **target language modification, tlm** (ORDER; ADD; MISS; CHOICE[11]; MERGE; SPLIT)

The increasingly fine-grained description of errors on levels one to three is specified in the AS. The first two levels are obligatory, while not for all error phenomena it might be possible to define level 3. However, all error tags must be used in the form defined in the AS; after piloting, no changes in the degree of specificity are allowed.

While a-c are three hierarchical levels of error description with growing level of specificity, d regards a way of error description that has become 'classical' (Corder 1974; discussion in Wisniewski's dissertation; Luedeling et al. 2012) and could be termed 'target language modification' (tlm, see Diaz-Negrillo et al. 2011?). There are 6 possibilities here: something might be in a place where it does not belong (ORDER), the learner might have chosen something wrongly (CH), added (ADD) or left out (MISS) something. The Falko team adds the category of MERGE, where parts of a linguistic structure that do not belong together are used in a merged form, and the opposite phenomenon, where parts belonging together are used separately (SPLIT). These aspects will be used in MERLIN, too.

---

[10] Full list. Further possibilities in EA2.
[11] Called CHANGE in Falko

Errors (Orthography and grammar) to be tagged:
(Precise descriptions and examples of the different tags can be found in the annotation scheme)

Grammar
Wrong word order (main clause/sub clause)
Errors in parts of speech
Missing/redundant obligatory complements
Errors in inflective forms (nouns, adjectives, simple and complex verbs (tense, aspect, mood, voice, use of auxiliary etc.)
Erroneous morphological aspects (case, number, gender)
Omission/addition of verbs
Errors in negation
Errors in reflexivity
Errors in the use of prepositions
Errors in the use of conjunctions
Errors in the use of clitics (ITA and CZE)
Orthography
Errors in the use of graphemes, accents and apostrophes
Erroneous abbreviations
Errors in punctuation
Errors in word splitting and merging

Structure resulting from Stage 1:

Tokenized learner text	…	…	…	…
Ctok	….	…	…	…
ZH1 (=TH1)	…corrected error1…	…	…corrected error2…	…
ZH1Diff	= CHA | INS | DEL | MOVS | MOVT | MERGE | SPLIT		= CHA | INS | DEL | MOVS | MOVT | MERGE | SPLIT	
ZH1spec	= 1 | 2 | 3 if applicable		= 1 | 2 | 3 if applicable	
EA1_lev1	**orth/gram**		**orth/gram**	
EA1_lev2	**…/…/…**		**…/…/…**	
EA1_lev3	**…/…/…**		**…/…/…**	

EA1_tlm	…/…/…		…/…/…	
Comment tier				

Table 2: Annotation levels available after Stage 1 (detail from Table 1)

Explanation of table 2:

…   =   filled fields

For an explanation of abbreviations, see Table 1.

From Table 2, you can see that TH1 has all fields filled with information, but not in every field there is the correction of an orthographic or grammatical error. If such an error is not found, TH1 is exactly the same as ctok.

You can also see that not in every field there is an error annotation. Only the phenomena that we have chosen as interesting, feasible, reliable and valid enough to be tagged in Merlin will be annotated (annotation scheme).

Please note: TH1 – EA1 consistency

Generally, where there is a change of ctok on TH1, there should also be an EA1 tag. The only exceptions are phenomena for which we did not include tags in the annotation scheme which represents a selection of relevant tags. That means that in single cases there might be a TH1 without a EA1 tag. This should be an exception; consistency between TH1 and EA1 is very important!

Please note: Incorrect TH1

Do not accept incorrect TH1, particularly if this would 'harm' the impression of the original language learner's texts. Correcting TH1 errors related to single tokens require different procedures than correcting errors that regard longer stretches of speech. Please refer to the Annex to the annotation structure guidelines (Vs4) for details.

V. 4 INTRODUCING STAGE 2: THE MERLIN CORE CORPUS

Extended target hypotheses (TH2) and error annotation 2 (EA2) will be carried through for a **core corpus of A2 and B2** texts. This will allow for meaningful comparisons between clearly distinct productions in terms of human ratings.

These levels are based on the **fair averages** that were arrived at in a Multi-Faceted Rasch analysis. They do not necessarily represent the level of the test the learner took, but they are a rating that has been statistically corrected for bias such as differing leniency/strictness.

For each language, roughly **200 texts** should receive TH2 and EA2. For **German and Czech**, the core corpus contains texts that have an A2 and B2 fair average. For **Italian**, there are N=209 texts available that received ratings corresponding to an A2 (no plus level). N=102 are chosen for the core corpus. Because only 2 texts were placed beyond the B2 threshold on the Rasch logit scale, apart from those the best N=99 texts on the logit scale were chosen for the corpus. The aim of the core corpus is to build two groups of clearly distinct learner productions. The more proficient group of texts contains B1 ratings (N=47), B1+ ratings (N=52) and B2 ratings (N=2). In total, the Italian core corpus aimed at contains **N= 203** texts.

The lists of files in the core corpus are contained in the **annotation progress sheets (Overview_core_corpus_stage_2)**.

As far as possible, the **reliability files** are the same ones as in stage 1. However, as there were some readjustments based on the calculation of the fair averages, many had to be newly assigned.

Detailed advice on how to carry out TH2 technically can be found in the Annex to these annotation structure guidelines.

V.5 STAGE 2: EXTENDED TARGET HYPOTHESIS (TH2)

After having described orthographic and grammatical errors in TH1 and EA1, we will write extended target hypotheses (TH2). These involves more subjectivity and difficulties of decision reliability, which is why they have been separated from the first level of target hypotheses just like in the Falko project. The aim of TH2 is to capture the perspective of **acceptability** of the learner text (not, like for TH1, its correctness). To this aim, the learner text should still be only minimally modified while at the same time its reconstruction should come close to what a native speaker utterance would look like.

This reconstruction regards semantic and lexical aspects, pragmatics, and sociolinguistics. Please do not correct the learner text if what you disagree with refers to an interpretation of the task that doesn't make sense to you or to an attitude towards the task you are not in line with.

In Falko, there is no full list of annotation rules to follow, as the phenomena are simply too numerous. However, we will use the existing Falko rules and also some MERLIN-specific rules, all of which are listed below. Furthermore, we will have to be aware of the phenomena we discussed during TH1 annotation as to be annotated on TH2 (see chapter 6: interactions between annotation layers).

TH2 will be based on a complete copy of the TH1 (created with help of the Falko Excel Add-in. In the end, all EA1 annotations will be reflected in TH1, and all EA2 annotations will be reflected in TH2. The opposite, however, is not necessarily true: There might be TH2 modifications that are needed to arrive at an acceptable version of the learner text and that are simply NOT part of our annotation scheme. In other words: in some cases there might not be an error tag in the annotation scheme that would be needed to describe a TH2 modification, as the number of tags in the scheme had to be limited.

Sometimes, corrections on TH2 level will imply changes of the morphosyntactic structure. Please note: those changes are **not** to be annotated (e.g. changing the verb in the main sentence requires a change of the mode in the subordinate clause; a changed verb requires the adaption of complements in TH2).

NO GRAMMAR AND ORTHOGRAPHY TAGS (from EA1) MUST BE ADDED FROM NOW ON!

Falko TH2 rules for all languages

1   Subclauses without introducing element → are not corrected, if a) they do not start with an embedded subclause themselves, b) the verb is in the final position or c) the verb normally appears with an introduced subclause

Please note: This rule to some of us wasn't very clear, so we contacted the Falko team and are now waiting for a response!

ctok: Ich meine, es lohnt sich nicht zu denken, dass Terror der einzige Ausgang ist.
ZH2: Ich meine, es lohnt sich nicht zu denken, dass Terror eine Lösung ist.
    (usb015_2006_10)

ctok: Mein Fazit lautet deshalb, wenn man im Leben Schwein hat, kann sich auch Kriminalität lohnen,
ZH2: Mein Fazit lautet deshalb, dass sich, wenn man im Leben Glück hat, auch Kriminalität lohnen kann. (cbs002_2007_10)

ctok: Deshalb finde ich nicht, dass man sagen kann, die Universitätsabschlüsse von einem geringeren Wert sind.
ZH2: Deshalb finde ich nicht, dass man sagen kann, dass die Universitätsabschlüsse von einem geringeren Wert sind (cbs003_2007_10)

ctok: Man soll auch anerkennen, ein Universitätsabschluss bereitet die Jugend auf die Welt vor, weil die meisten einflussreichen Menschen der Welt schön einen Abschluss haben.
ZH2: Man sollte auch anerkennen, dass ein Universitätsabschluss die Jugend auf die Welt vorbereitet, weil die meisten einflussreichen Menschen der Welt schon einen Abschluss haben. (fu120_2006_10a)

ctok: Beim Abschluss erkennt der Student, er kann schwierige Ziele erreichen und sein Selbstwert nimmt zu.
ZH2: Beim Abschluss erkennt der Student, dass er schwierige Ziele erreichen kann und sein Selbstwert nimmt zu. (fkb055_2008_08)

6   Definiteness: if there is a context-dependent error in the expression of definiteness of a referent, please correct this

ctok: Das ist wahr, dass es nicht so für die Frauen in allen Ländern ist, aber das ist keine Folge **von** Feminismus.

ZH2: Es ist wahr, dass es nicht für die Frauen in allen Ländern so ist, aber das ist keine Folge **des** Feminismus. (fkb031_2008_07)

ctok: Ich finde, dass **eine finanzielle Entlohnung eines Menschen** dem Beitrag entsprechen soll, den er / sie für das Unternehmen leistet.

ZH2: Ich finde, dass **die finanzielle Entlohnung eines Menschen** dem Beitrag entsprechen sollte, den er / sie für ein Unternehmen leistet. (fkb010_2008_07)

1   Tempus and Mode are adapted to the context

ctok: Wie die Berliner haben von Barrack Obama gerade gestern gehört, unsere Sozietät **glaube**, und muss glauben bleiben, der Idee der Ebenheit und der Gleichkeit aller Personen, Mann oder Frau.

ZH2: Wie die Berliner von Barrack Obama gerade gestern gehört **haben**, **glaubt** unsere Gesellschaft und muss weiter glauben an die Idee der Gleichheit aller Personen, ob Mann oder Frau.
(fkb031_2008_07)

ctok: Wegen des Obenstehendes ist es also nicht in Ordnung zu äußern, dass die Universitätsabschlüsse von geringem Wert **sind**.

ZH2: Deshalb ist es also nicht in Ordnung zu sagen, dass die Universitätsabschlüsse von geringem Wert **seien**. (cbs007_2006_09.)

2   Expression of mode and aspect → uncommon expressions of mode and aspect are corrected

ZH2: Wie die Berliner von Barrack Obama gerade gestern gehört haben, glaubt unsere Gesellschaft und muss **weiter** glauben an die Idee der Gleichheit aller Personen, ob Mann oder Frau.
(fkb031_2008_07)

ctok: Wie die Berliner haben von Barrack Obama gerade gestern gehört, unsere Sozietät glaube, und muss glauben **bleiben**, der Idee der Ebenheit und der Gleichkeit aller Personen. Mann oder Frau.

3   Untypical and uncommon use of modal verbs→ please correct

ctok: Obwohl die Lehren so wichtig ist, **können die Lehrer nicht gut verdienen**.

ZH2: Obwohl die Lehrer so wichtig sind, **verdienen** sie nicht gut. (trk008_2006_05)

4   connectors → if connectors express an unlogical relation in the text, replace them (see MERLIN-specific rule for coherence/cohesion as well; EA2 tag: C_Con_accur]

ctok: Weil für meisten Menschen das Geld als die beste Entlohung gilt, wollen die andere andere Entlohungen, schätzen die andere Werte, glauben an die unsiegbare Macht des Gelds nicht.

ZH2: **Während** für die meisten Menschen das Geld als die beste Entlohnung gilt, wollen andere eine andere Entlohnung, schätzen andere Werte und glauben nicht an die unbesiegbare Macht des Gelds. (fu127_2006_10c)

ctok: müssen uns auch fragen, ob die These des oberliegenden Titels heutzutage wirklich erreichbar sei, **weil** in diesem politischen Klima, wer hat die Macht, solche Entscheidungen über Lohne und entsprechende gesellschaftliche Beitrage eigentlich zu treffen?

ZH2: Wir müssen uns auch fragen, ob die These des obigen Titels heutzutage wirklich erreichbar ist. **Denn** wer hat in diesem politischen Klima eigentlich die Macht, solche Entscheidungen über Löhne und entsprechende gesellschaftliche Beiträge zu treffen? (hu006_2006_09)

5   Missing objects → in case of SEMANTICALLY missing objects, please add them (syntactically missing objects have been inserted in TH1 already)

ctok: Nach seine Universitätsabschluss kennt er wahrscheindlich jeder Wirkung von eine besondere Maschine aber er hat noch nie früher die Gelegenheit physisch mit dieser Maschine um zu gehen und sich genauer zu beschauen.

ZH2: Nach seinem Universitätsabschluss kennt er wahrscheinlich jede Wirkung einer besonderen Maschine, aber er hatte früher noch nie die Gelegenheit, physisch mit dieser Maschine umzugehen und sie sich genauer zu beschauen. (sa001 2006 09)

6   Numerus → if a NUMERUS is hardly plausible, correct

ctok: . Unsere Krankenschwester und andere Mitarbeiter des öffenlichen Unternehmens haben ihre Meinungen hierzu geäussert.

ZH2: . Unsere Krankenschwestern und andere Mitarbeiter in öffentlichen Unternehmen haben ihre Meinung hierzu geäußert. (fkb010_2008_07.)

7   Reference inside a matrix sentence → if inside a sentence there is wrong reference with anaphors and cataphors, please correct [EA2 tag: C_coh_ref]

ctok: Wenn man denkt, dass Kriminalität auszahlt, sollten sie über die Folgen auch denken.

ZH2: Wenn man denkt, dass Kriminalität sich auszahlt, sollte man auch über die Folgen nachdenken. (BNG2-2010-11-116)

8   Reference across matrix sentences → will only be corrected if more than one version is possible in one matrix sentence [EA2 tag: C_coh_ref]

ctok: Es kann passieren dass die Leute verändern sich und auch seine Meinungen im Knast.
Aber bekommen sie noch eine andere Chance vom Leben und vom Bekannten?
Wird **ihm** jemand hilfen, oder werden sie kein Chance haben sondern nur eins - noch einmal ausprobieren ein besserer und kluger Kriminalist zu werden?

ZH1: Es kann passieren, dass die Leute im Knast sich und auch ihre Meinungen verändern.
Aber bekommen sie noch eine andere Chance vom Leben und von Bekannten?
Wird **ihnen** jemand helfen oder werden sie keine Chance haben, sondern nur eines : noch einmal ausprobieren, ein besserer und klügerer Krimineller zu werden? (fkb005_2007_09)

B   SEMANTICS

9   Replace verb, German Funktionsverbgefüge: reconstruct around the nominal nucleus

ctok: Vielleicht auch manche Frauen würden gerne diese Meinung vertreten, vor allem die, die anspruchsvolle Arbeit ausüben müssen um Geld zu **gewinnen**.

ZH2: Vielleicht würden auch manche Frauen gerne diese Meinung vertreten, vor allem die, die anspruchsvolle Arbeit ausüben müssen, um Geld zu **verdienen**.
(fkb030_2008_07)

ctok: **Oder Vorteile besser zu sagen**. Wir können, sogar müssen, das selbe wie die Männer schaffen. (fkb034_2008_07)

ZH2: Oder, **um eher die Vorteile zu nennen**: Wir können und müssen sogar, dasselbe schaffen wie die Männer.(fkb034 2008 07)

10   non-standard reduction of particle verbs → correct

ctok: Und ich bin der Meinung, dass es dann schwerer ist da wieder rauszukommen.

ZH2: Und ich bin der Meinung, dass es dann schwerer ist, da wieder herauszukommen. (dew18_2007_09)

11 formulaic sequences → if there is an uncommon formulaic sequence in the learner text for which there is a common one in the target language, please correct. This refers to all types of formulaic sequences, including collocations. Please have an extra eye on the annotation scheme for this.

ctok: In der dänischen Gesellschaft <u>werden</u> diese Ausbildungen sehr **nachgefragt**, weil es zur Zeit ein Mangel an diesen Fachkräften besteht.

ZH2: In der dänischen Gesellschaft <u>sind</u> diese Ausbildungen sehr **gefragt**, weil zur Zeit ein Mangel an diesen Fachkräften besteht. (fkb049_2008_08)

12 Modal particles → if a common modal particle is missing in a construction, please correct

ctok: <u>Sei die im Titel benannte These provozierend</u>, und auf ersten Blick völlig unbegründet scheinen, auf jeden Fall ist sie eine Diskussion wert.

ZH2: <u>Sei die im Titel benannte These **auch** provozierend</u> und mag sie auf den ersten Blick völlig unbegründet scheinen, auf jeden Fall ist sie eine Diskussion wert. (fkb058_2008_08)

13 Lexical choices → if a learner uses an uncommon word/formulaic sequence instead of a common one to denote something, and if it is clear that he/she did not do that on purpose/by choice, the word/formulaic sequence will be corrected. [EA2 tags: V_lexgrammerr_incompr, V_semdenot_word/fs, V_semconn_att_word/fs, V_sem_imprec, V_form_wordform, V_form_word/fs_nonexist, V_form_FS, V_form_FS_disrupt)

ctok: Es gibt gewaltsam Kriminalität wie Mord, Kriminalität gegen **humanität**

ZH2: Es gibt gewaltsame Kriminalität wie Mord, Kriminalität gegen die **Menschlichkeit** (kne20_2006_07)

ctok: Die Errungenschaften der deutschen Philosophen, der deutschen Historiker des 19. Jhs bewertet man sich sehr hoch in der **Weltwissenschaft**.

ZH2: Die Errungenschaften der deutschen Philosophen und der deutschen Historiker des 19. Jhs. bewertet man in der **globalen Wissenschaftsgemeinde** sehr hoch. (fkb051_2008_08)

14 Adapt references to context across the sentence limit [EA2 tag: C_coh_ref]

ctok: Sie hatten keine Lust mehr sich nur <u>um **eigene Kinder**</u> zu kümmern, Essen zu ko-
chen, Wäsche zu waschen, kurz zusammengefasst: zu Hause wie in einem Käfig
zu sitzen.

ZH2: Sie hatten keine Lust mehr, sich nur <u>um **die eigenen Kinder**</u> zu kümmern, Essen
zu kochen, Wäsche zu waschen, kurz gesagt, zu Hause wie in einem Käfig zu sit-
zen. (fkb030_2008_07)

C   STYLE

15   direct speech will only be corrected if it is wrong

ctok: Wozu nutzt es, wenn einer beim Bewerbungsgespräch sagt: "Ja, ich habe die The-
orien von Marx, Weber und Durkheim im Griff und ich bin ein netter und pünkt-
licher Mensch."?

ZH2: Wozu nutzt es, wenn einer beim Bewerbungsgespräch sagt: "Ja, ich habe die The-
orien von Marx, Weber und Durkheim im Griff und ich bin ein netter und pünkt-
licher Mensch."? (cbs011_2006_09)

ctok: Was soll die Lösung sein? Sicherlich nicht die Satz "<u>Universitätsabschlüsse</u> sind
von geringem Wert".

ZH2: Was soll die Lösung sein? Sicherlich nicht der Satz: "<u>Universitätsabschlüsse sind
von geringem Wert." (fk004_2006_08.)</u>

16   Introductions in a "journalistic style" are not corrected

ctok: <u>Das Unterschied:</u> das Praktikum muss völlig anhand des Studenten ausgehen und
wird nur bis zum 50 Prozent als Teil des Studiums anerkannt.

ZH2: <u>Der Unterschied:</u> Das Praktikum muss völlig von dem Studenten ausgehen und
wird nur bis zu 50 Prozent als Teil des Studiums anerkannt. (cbs005_2006_09)

17   relative pronouns → are replaced if they do not correspond to the standard

ctok: Die Schule, <u>wo man streng kontrolliert wird</u>, ist schon in der Vergangenheit ge-
blieben.

ZH2: Die Schule, <u>in der man streng kontrolliert wurde</u>, ist schon Vergangenheit.
(fkb044_2008_08)

**18   vocabulary → colloquial forms are replaced by more appropriate ones. If you are in doubt,
please consult a dictionary**

ctok: Wegen der methodischen / theoretischen Vorgangsweise **kriegt** eine Wissen-
schaftliche Arbeit auch eine besondere Form.

ZH2: Wegen der methodischen bzw. theoretischen Vorgehensweise **bekommt** eine wis-
senschaftliche Arbeit auch eine besondere Form. (cbs013_2006_09)

D   INFORMATION STRUCTURE

1   Information structure: status of information → if a nominal phrase refers to a referent that has been mentioned already in the same sentence, it is pronominalized

ctok: Die Unis sollten auch besser <u>den Studenten</u> erklären, wofür <u>sie</u> etwas lernen. Das,
was **man** lernt sollte **der Student** zu einer bestimmten Anwendungssituation ver-
binden können.

ZH2: Die Unis sollten <u>den Studenten</u> auch besser erklären, wozu <u>sie</u> etwas lernen. Das,
was <u>sie</u> lernen, sollten <u>sie</u> mit einer bestimmten Anwendungssituation verbinden
können. (cbs006_2007_10)

ctok: Die meisten Frauen möchten, damit sich <u>die Männer</u> <u>mit ihnen</u> manchmal wie
Gentlemen umgehen, **ihnen** Hilfe leisten.

ZH2: Die meisten Frauen möchten, dass die Männer ihnen wie Gentlemen Hilfe leisten.
(fk019_2006_07)

ctok: Der Mann spielte eine zentrale Rolle in der Gesellschaft. **Der Mann** war nicht der
Frau untertan

ZH2: Der Mann spielte eine zentrale Rolle in der Gesellschaft. **Er** war der Frau nicht
untertan (fk024_2006_07)

2   information structure: focus → a focus the structure of which is not appropriate for the context will be adapted

ctok: **Die besonders schwere** Lage war in Deutschland.

ZH2: Besonders schwer war die Lage in Deutschland.
(fkb051_2008_08)

ctok: Es gibt aber Häuser, wo die Mutter zu Hause nur faulenzt.
<u>Es ist so</u> beispielsweise in der Familie meines Freundes

ZH2: Es gibt aber Haushalte, in denen die Mutter zu Hause nur faulenzt.
So ist <u>es</u> beispielsweise in der Familie meines Freundes (fk015_2006_07)

3 focus particles → the position of the focus particles will be adapted to the intended scopus (?)

ctok: Dies Erkentniss ist <u>auch</u> <u>wesentlich</u> <u>in finanziellen Berufen</u>, und in Berufen, in den man muss Argumenten machen und schützen, zum Beispiel politische Berufen.

ZH2: Diese Erkenntnis ist <u>auch in der Finanzbranche</u> <u>wesentlich</u> und in Berufen, in denen man Argumente bringen und unterstützen muss, zum Beispiel in der Politik. (fk013_2006_08.)

4 information structure: topic → expressions of topic precede expressions of focus if the focus can not be regarded as topicalized in the context

ctok: Z. B. Leibniz hatte keine Möglichkeit <u>an der Leipziger Universität</u> zu promovieren, <u>denn</u> **die grosse Korruption** <u>herrschte dort</u>, die Professoren halteten die schlechten Vorlesungen, nicht immer kamen zu den Vorlesungen u s. w.

ZH2: Z. B. hatte Leibniz keine Möglichkeit, <u>an der Leipziger Universität</u> zu promovieren, <u>denn</u> <u>dort herrschte</u> **große Korruption**, die Professoren hielten schlechte Vorlesungen, sie kamen nicht immer zu den Vorlesungen usw. (fkb051_2008_08)

Falko TH2 rules regarding GER/CZ

1 GER/CZ: Declarative sentences with verb in first position without subject → insert subject

ctok: Sie stossen auf Situationen wo sie nuicht genau wissen was sie tun sollten. **Fühlen sich etwas orientirungslos.**

ZH2: Sie stoßen auf Situationen, in denen sie nicht genau wissen, was sie tun sollten. **Sie fühlen sich etwas orientierungslos.** (cbs006 2007 10)

2 GER: "Nachfeld" → long prepositional phrases, comparative clauses, relative clauses and appositions will go into the "Nachfeld"

ctok: Oder Vorteile besser zu sagen. Wir können, sogar müssen, <u>das selbe</u> **wie die Männer** schaffen. (fkb034_2008_07)

ZH2: Oder, um eher die Vorteile zu nennen: Wir können und müssen sogar, <u>dasselbe</u> schaffen **wie die Männer**.(fkb034 2008 07)

3 GER: "dass"-sentences and infinitives with "zu" → an object clause in which the subject takes up a referent from the main clause is constructed with an infinitive structure with "zu"

> ctok: Sie suchen nach der Lösung und anklagen die Universitäten, <u>dass die keine gute</u> <u>Spezialisten</u> **bereiten können.**
>
> ZH2: Sie suchen nach der Lösung und klagen die Universitäten an, <u>keine guten Spezia-</u><u>listen</u> **auszubilden.** (fkb044_2008_08)

Please note:

Especially for Czech, but also for Italian, please inside your team explicitly agree upon a language variety that you want to underlie as acceptable before starting the annotations!!!

MERLIN-specific TH2 rules

These rules derive from the necessity to bring TH2 and EA2 in line with each other. The definitions of the EA2 tags are given in brackets behind the explanation of the TH2 correction rules in order to make the link between TH2 and EA2 more explicit. The tags will be applied only later, of course, when TH2 is finished. Also, there are more EA2 tags than appear on this list (all existence tags).

Only in cases where Falko rules are not explicit enough, an extra MERLIN rule is applied.

SOCIOLINGUISTIC APPROPRIATENESS

1. If a word/formulaic sequence is not appropriate in terms of **formality** required by the task (official/unofficial, personal/impersonal, colloquial language/standard language), please correct. This refers to the diastratic and diaphasic variation of language (register/style) [EA2 tag: S_form_gen]

2. If there is an inappropriate **addressing** in the learner text, please correct this (Du/Sie, tu/Lei, Czech equivalent). [EA2 tag: S_form_addr]

3. If there is a lack of politeness resulting in **overly direct language forms**, please correct this, too [EA2 tag: S_pol_dir].

4. ITA: Please correct inappropriate **double pronouns** which belong to spoken language and shouldn't be found in written productions (e.g.: "A me mi piace andare al cinema"). [EA2 tag: S_Var_duppron]

5. ITA: Please correct **the "che polivalente"** which introduces subordinate clauses that would need more precise conjunctions, e.g. "ho sentito cose che non avevo pensato", or "vieni qui che ti voglio dare qualcosa". [EA2 tag: S_Var_che]

6   GER: please correct main clause sentence structure after "**weil**" [EA2 tag: S_Var_woweil]

7   GER: Please correct **wrongly used discourse particles** halt, ja, eben, denn, doch, aber, auch, ruhig, nur, mal, schon, überhaupt, vielleicht, wohl (please note: this is not to be mistaken with adverbs). [EA2 tag: S_Var_partik_0]

COHERENCE/COHESION

1   Correct wrong connectors. Insert missing connectors. Comply to the FALKO rules (see above). [EA2 tags: C_Con_accur]

2   If you feel that there is a "Content jumps", i.e. a sudden change of topic without any linguistic means of introducing it that interferes with cohesion, please write a TH2. [C_con_jump]

VOCABULARY

1   correct inappropriate but grammatically acceptable choice of positive/comparative/superlative (*FAQ-document (Italian, German, Czech)*

2   correct inappropriate but grammatically acceptable choice of interrogative pronouns / use of adverbs *FAQ-document (Italian, German, Czech)*

3   for ITA, correct: confusion of interrogative and demonstrative adjectives *FAQ-document (Italian, German, Czech)*

GRAMMAR

1   TH2 will be written also for grammar errors that can be observed in context only and thus have not been corrected in TH1. No EA1 tags will be used in the EA2 process, though!! The following phenomena have been discussed: inter-sentential mood and tense conflicts (see Falko rules) (→ reference: FAQ-document: "EA1-corrections-onTH2")

2   wrong choice of **definite/indefinite article** to be observed in **context** only (→ reference: FAQ-document: "EA1-corrections-onTH2")

3   wrong choice of a **preposition** to be observed in **context** only (→ reference: FAQ-document: *"EA1-corrections-onTH2")*

V.6 STAGE 2: ANNOTATION OF LEXICAL, SOCIOLINGUISTIC, PRAGMATIC and INTELLIGIBILITY ASPECTS IN MERLIN

In the last annotation step, phenomena regarding the **intelligibility** of the learner texts, their **lexical** characteristics as well as **sociolinguistic** and **pragmatic** aspects (coherence/cohesion, speech act REQUEST, politeness) are taken into consideration for the MERLIN core corpus. This means that for productions that already received TH2, EA2 is annotated as well. EA2 is based on TH2.

Technically, EA2 is similar to EA1; MMAX2 is used to annotate EA2.

Descriptive existence tags

A special characteristics of EA2 are the "**existence tags"**. In Merlin, not only learner errors are annotated. From a more positive and constructive perspective, we also ask 'what's there' (for example, we tag opening and closing formulae). These tags are used to tag phenomena in learner language no matter if they are realized correctly or not, for example the speech act REQUEST or formulaic sequences (FS); the approach is descriptive, not prescriptive/normative

In the annotation scheme, tags that do not focus on errors are marked in grey. All occurring instances must be **annotated regardless of their correctness.**

EA2 existence tags

V_FS_colloc	formulaic sequence: collocation
V_FS_colloc_compeq	formulaic sequence: compound equivalent (ITA)
V_FS_idiom	formulaic sequence: idiom
V_FS_commphras	formulaic sequence: communicative phraseologism
C_Coh_txtstruct	metacommunicative device
S_Txt_grfw	salutations/complimentary closes
S_Txt_opcl	opening/closing formulae
S_Var_clit	ITA: lexicalised clitics (verbi procomplementari)
S_Var_synstr	ITA: marked syntactic structures
S_Var_partik	GER: modal particles
P_Pol_dir	politeness - overly direct language form
P_Request_direct	direct REQUEST

EA2 error tags

G_Intelltxt	intelligibility of text
G_Intells	intelligibility of sentence
V_Sequence_lexgrammerr_incompr	incomprehensible sequence caused by accumulation of lexical/grammatical error(s)

V_Word_form_nonexist	non-existing form (word)
V_FS_form_nonexist	non-existing form (formulaic sequence)
V_FS_form_incompr	formulaic sequence: limited intelligibility
V_Word_semdenot	semantic error: denotation (word)
V_FS_semdenot	semantic error: denotation (formulaic sequence)
V_Word_semcon_att	semantic error: connotation (attitude), (word)
V_FS_semcon_att	semantic error: connotation (attitude), (formulaic sequence)
V_Word_semimprec	semantic error: precision (word)
V_FS_semimprec	semantic error - precision (formulaic sequence)
V_Wordform_deriv	word formation error: derivation
V_Wordform_comp	word formation error: composition
V_FS_form	formulaic sequence: form error
C_Con_accur	connector accuracy
C_Coh_jump	content jumps
C_Coh_ref	Reference
S_Form_gen	inappropriate style (formality)
S_Form_addr	inappropriate addressing (formality)
S_Var_duppron	ITA: personal pronoun redundancy
S_Var_che	ITA: 'che polivalente'
S_Var_woweil	GER: main clause word order after 'weil'
P_Request_indirect	indirect REQUEST

Detailed definitions and examples are given in the annotation scheme. Reference works (e.g., connector lists, dictionaries and grammars) are also supplied in the annotation scheme.

Reliability of annotation

For EA2, we cannot expect a perfect agreement between annotators because many things are subjective. However, **reliability checks** are carried through for the same files that also underwent TH2 reliability procedures.

Further explanations on EA2

In the **Annex** to these Guideline, you find instructions about …
- how to treat **TH2 errors** you encounter when annotating EA2
- how to treat **EA1 errors** you encounter when annotating EA2
- where to **save / commit** EA2-annotated files
- **reliability** procedures

Structure resulting from Stage 2

Tokenized learner text	....	...	...	...
Ctok	....	...	...	...
ZH1 (TH1)	...corrected error1...	...	...corrected error2...	...
ZH1Diff	= CHA | INS | DEL | MOVS | MOVT | MERGE | SPLIT		= CHA | INS | DEL | MOVS | MOVT | MERGE | SPLIT	
ZH1spec	= 1 | 2 | 3 if applicable		= 1 | 2 | 3 if applicable	
EA1_lev1	orth/gram		orth/gram	
EA1_lev2	.../.../...		.../.../...	
EA1_lev3	.../.../...		.../.../...	
EA1_tlm	.../.../...		.../.../...	
ZH2 (TH2)		...corrected phenomenon1		...corrected phenomenon2....
ZH2Diff		= CHA | INS | DEL | MOVS | MOVT | MERGE | SPLIT		= CHA | INS | DEL | MOVS | MOVT | MERGE | SPLIT
EA2_lev1		soc/lex/coh		soc/lex/coh
EA2_lev2		e.g. formality		e.g. reference error
EA2_lev3		e.g. spec.: inadequate addressing		
EA2_tlm		If applicable. E.g. CHOICE		
Comment tier				

Table 3: Annotation levels available after Stage 2 (detail from Table 1)

Explanation of table 3:

... = **filled fields**

For an explanation of abbreviations, see Table 1.

Stage 2 annotation is marked in blue. The grey cells are part of Stage 1 (III.2.2 and III.2.3).

From Table 3, you can see that TH2 is only filled with information if an error that we decided to tag (Annotation Scheme) can be found. That means that **in Stage 2, if there is a target hypothesis, there is also an error annotation.**

V.7 Further annotation levels(s): ctok, ZH1Diff, ZH1spec

Ctok

Although we check tokenization and transcription errors before the annotation starts, it is possible that there will still be errors that will be detected only during annotation. Therefore, the level all other annotations are based upon is ctok, where corrections of transcription errors or tokenization are still possible. Ctok basically constitutes the learner text that has been transcribed.

Corresponds to ctok-level in FALKO Essay L2 corpus. The same abbreviation will be used inside MERLIN. This step is important for word counts, for example, where we do not want things like '20 . Jahrhundert' to be counted as 3 words.

ZH1Diff / ZH2Diff

While creating the TH1 (and also TH2) the automatic Diffmark function of the Falko Add-in is used to create surface error markers (for changes, deletions, insertions and movements). Thus the two tiers ZH1Diff and ZH2Diff illustrate all differences between the learner's text (ctok) and TH1/TH2, i.e. the changes that have been made by the annotator.

ZH1spec

Level for marking speculative hypotheses (1=uncertain, but inferable; 2=TH1 is a wild guess, 3= no TH1 possible as partly not comprehensible). By using this level those parts of learner productions are marked where deviating hypotheses are most probable, e.g. incomprehensible or very uncertain learner sentences and missing words. They thus can be excluded from further calculations, if necessary.

Comment tier

If there are unclear decisions or things alike, the annotators can give a comment in the comment level.
→ see ZH1spec

VI   Interactions between annotation layers

When carrying out Stage 1 annotations, it was not always easy to decide on which level of annotations to annotate single phenomena. These phenomena were discussed in the language groups in email and skype discussions. They are documented in the FAQ document, sheet "EA1 corrections on TH2/EA2" ((https://docs.google.com/spreadsheet/ccc?key=0AouRwmLOZeQudEw2ZmJ0dGtzQm5Nc0c1U3NjTVd mOHc&disco=AAAAAFYonic#gid=11).

It is important to understand the relationship between the annotation layers in MERLIN. This graphic is meant to illustrate them:

Learner text

TH1

TH2

EA1

EA2

TH1 & EA1: Completely reflected vice versa, aim: consisten

EA1 & EA2: complementary, independent

TH2 integrates TH1

a) every EA2 tag is reflected in a modification of the learner langu

b) many, but not all modifications of the learner language on TH
reflected in an EA2 tag

EA2 & TH1: independent
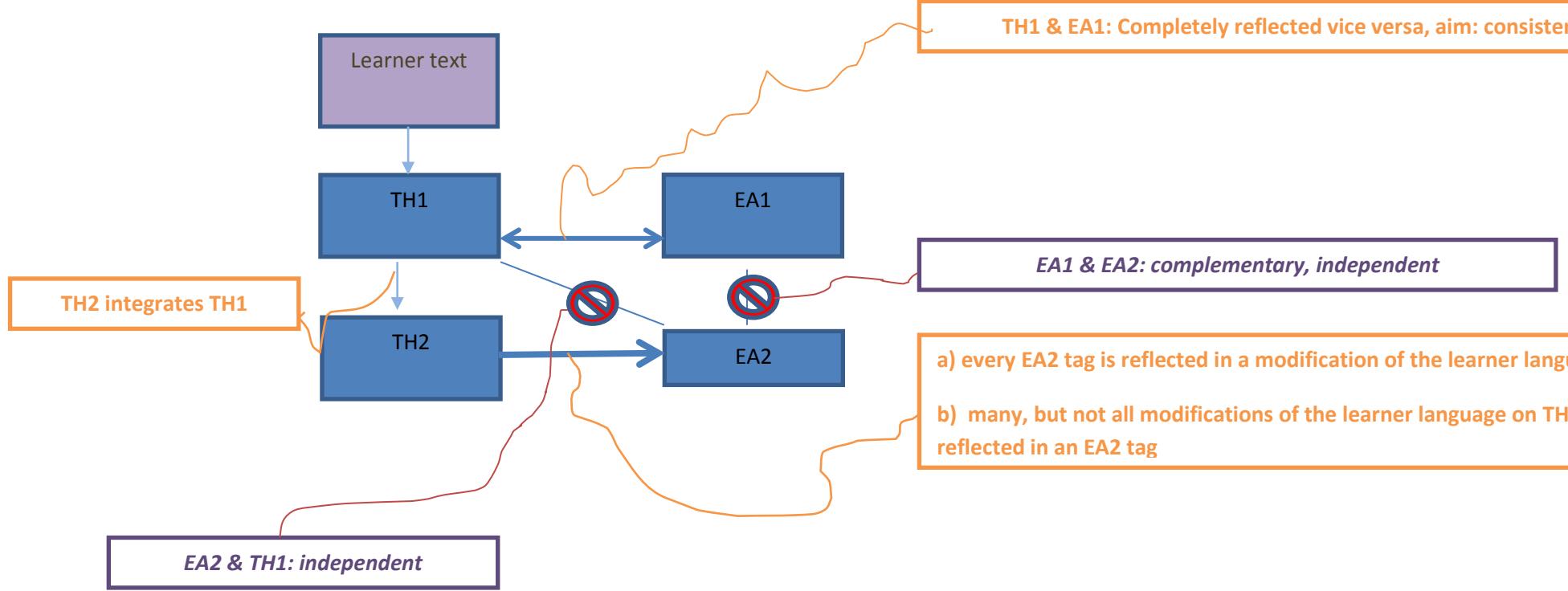
Graphic xyz: Annotation layers in MERLIN

47

TH1 & TH2

TH2 includes TH1, it represents an extension of the minimal first target hypothesis. Phenomena that are to be taken into consideration when writing TH2 are, for example:
- phenomena that can be observed in **context** only, like
    o inter-sentential mood / tense conflict
    o wrong choice of definite/indefinite article
    o usage of connectors not appropriate in context; eps.: GER: relative pronoun „wo"
    o wrong choice of a preposition to be observed in context only
- inappropriate informational structure
- inappropriate choices of comparatives/superlatives
- inappropriate choice of interrogative pronouns or adverbs
- …

For examples, please refer to the FAQ document.

Many of these TH2 changes will not be reflected in EA2, because there is no correspondent EA2 tag. In the list above, an exception is the connector example (C_Conn_accur). Also, no EA1 tags will be applied when carrying out EA2.

TH2 & EA2 consistency

Please mind: While in STAGE 1, for almost all corrections carried out on level TH1, there also was a tag on EA1, this will not be the case in STAGE 2 because there are simply too many possible modifications in order to arrive at an appropriate text form (TH2) to be specified in the Annotation Scheme as part of EA2. Thus, it is well possible that you correct a learner language on TH2 without later being able to add an EA2 tag (e.g.: information structure).

The other way round, if there is an EA2 tag, be sure that TH2 reflects this, as this should always be possible.

EA1 & EA2

What is very important, because we changed this during the development of our corpus: **Please keep in mind: Even if a TH2 modification refers to grammar, no EA1 tags will be added in EA2!**

When writing EA1, we specified a number of phenomena that were not annotated in EA1, but will have to be annotated in EA2. Please refer to the FAQ document and the Annotation Progress Sheet for details.

Phenomena are, for example:
- cases that were hard to classify as pertaining to orthography (EA1) or lexis (EA2)
- cases that were hard to classify as pertaining to grammar (EA1) or lexis (EA2)
- foreign words (EA2)
- …

VII Annex

Annex I: Guidelines for annotation of EA1 – TH2 – EA2 (Annex to the Annotation structure_Vx.docx in: moodle / WP5 / Error annotation tools and guidelines and on the svn → reports)