Jochen Rehbein (2008) [preliminary draft]

Transcribing spoken language with *EXMARaLDA / Partitur-Editor*[1][2]

– step by step instructions based on an example of spoken Turkish

I. *Download Exmaralda from*

http://www.exmaralda.org/index.html
> http://www.exmaralda.org/downloads.html
>Partitur-Editor (choose version appropriate for your system)

II. *Open Partitur-Editor* (precondition for version 1.3.4 s Java 2)

III. *New file*

Menue > File > (window "Sichern unter":) Save basic transcription as „(give name here)" as file on hard disk (> Dateiformat: Extensible Markup Language) > Speichern (= save) > check, if the transcript file is saved under the name you gave in the folder you chose

IV. *Enter metainformations about the new file*

Menue > File > Meta-Information > Edit meta information

a) Fixed attributes:
   > project name: xxx; Transcription: xxx; referenced media file: enter names for digital recordings only (audio data)

b) User defined attributes (list of attributes): > Add attribute (click) > replace under „Attribute" New1 by date of recording> fill in under „value", e.g. 150508 (for 05/15/2008)
   [[enter as much information as possible, here; e.g. on:
   • childrens' age; • family/experiment of children; • situation of communication; • bilingual/monolingual (= "Versuchsgruppe"); institution where the recording took place; • recording person; • type of recording apparatus used; • number of media; • beginning/end date of transcription; • proportion of recorded time related to time used for transcription = transcription proportion"); • time used for transcription control; • person who controlled transcription; • translator; • time used for translation; • apparatus used for transcription etc.]

   > Complete all entries > ok
   ((do not use the close function)) ((If there appears a further dialogue, nevertheless, click ok))

V. *Enter meta-information about speakers*

Menue > File > Speakertable > *Speakers* ((first speaker in the topmost line in the transcript; order of speakers can be changed every time after having entered speakers' abbreviations:)) SPKO [X]

---

[1] For a detailed presentation of the complete *EXMARaLDA program* including reasoning of the moduls and operations, s. the Manual of Thomas Schmidt 2006.

[2] A note on translation: German '**Partitur**', Dutch ' **Partituur'** is Engl. ***score***,; Gm. **'Partiturfläche'**, Dt. **'Partituurvlak'** is En. ***score area***, Gm. **'Partiturklammer'**, Dt. '**Partituurhaak'** is En. ***brace*** (s. Ehlich, Mackenzie, Rehbein,Thielmann & ten Thije 2006**). –** In the following, German **'Partitur'** is used as a technical term in the sense of Engl. ***score,*** as in *'Partitur-Editor'.* The term traces back to Ehlich & Rehbein 1976.

a) Speaker properties: > Abbreviation: enter abbreviation for speaker (speaker sigle), e.g. Funda > sex: female
b) Languages (click in the list of language names), e.g. TRK (for Turkish) > add languages used > abbreviation of the language(s) used appears under "Languages"
c) Under "User defined attributes" > Add attribute > New1 > double click > enter e.g. age; under Value > double click same line as the corresponding Attribute > e.g. 25; full name> e.g. Funda

Add next speaker
Go via closing "Edit speakertable" for the first speaker and open a new "Edit speakertable" window> Menue > File > Speakertable and so on
OR
directly go within opened window "Edit speakertable" to next speaker by clicking "Add speaker" > [SPK1] > e.g. Esra > sex: female > Add attribute > New1 e.g. age> Value, e.g. 25 > Add attribute > Attribute: e.g. full name > Value: e.g. Esra

Now, all speakers are listed with their sigles under "Speakers" in the left window, e.g.
SPK0 [Funda]
SPK1 [Esra]
SPK2 [Zehra]

ok ((if you like you may cancel)) > in the transcript file opened, speaker Funda [v] only is indicated; other speakers will be indicated, if you continue with
> Menue > Add tier > window opens "New tier" > SPK1 [Esra], Type: T(ranscription), Category: v > ok
AND
> Menue > Add tier > window opens "New tier" [instead: apple + A **OR** button **+** (in the middle)]> SPK1 [Zehra], Type: T(ranscription), Category: v > ok

((if there is only one speaker or the last one:))
> close window "Edit speakertable" > ok > now, all speakers provided with an abbreviation [or logogramme] are attributed a separate line [= **trace**] in the opened Exmaralda file.

VI. *Description of constellation; meta-information on constellation; open field of comment [= "Kommentarfeld"]*

> Menue > File > Meta information... > edit meta information > (lowest window:) Comment > put in here meta-information on constellation > ok
(If there appears a further dialogue, nevertheless, click ok)

((General remark: If v should not appear in the file, then erase it)) > Menue > tier > Edit tier properties > Display: > cancel tickmark in square "Auto" > erase v after "Display"; replace it by e.g. [Eng] (but do not touch the entry "v" after "Category") > ok

VII. *Formatting tiers (style, script, point etc.)(G I.) and tier labels (H)*

In Exmaralda, you have a choice between 3 types of tiers:

1. ***Type T(ranscription)*** *for the verbal spoken utterances (the data of this tier will be entered into the data base)*
2. ***Type A(nnotation)*** *for translations and morphological transliterations; also for intonation and prosody (s. I below)*
3. ***Type D(escription)*** *for comments, nonverbal communication; type D is used as*
   a. *speakerspecified tier [with the label nv]* for comments, for non verbal communication (=nvc) etc.;
   b. *speakerunspecified (nonverbal) tier [with the nn]* for comments of those acoustic phenomena which cannot to be attributed to a certain speaker, e.g. a slammed, car noise, unidentifiable voices in the background etc.

Now follow the instructions how to create the correspondend tier:

I. Formatting (and creating) tiers

1) Verbal tier: Arial Unicode MS, 16 point, Type T(ranscription)

> activate a speaker's verbal tier by mouse click into the speaker's tier label field > Menue > Format > Format tier (or command + F) > Edit tier format
> Font: Arial, Font style: Plain, Font size: 16 [you may also use properties *Text colour, Background colour, Alignment*] > ok
((repeat procedure for all speakers, here e.g. Esra and Zehra:))

2) Translation tier, Type: A(nnotation)

a. *Create Translation tier*
   > Menue > Tier > Add tier (command + A; or click icon) > Speaker: [no speaker], Type: A(nnotation) > ok (a new tier [v] appears in partitur) > Activate new tier [v] by mouse click into the name field > Menue > Tier > Edit tier properties > speaker: [no speaker] > Type: A(nnotation) > Category: v > Display: > cancel tickmark in square "Auto" and erase v after "Display"; replace it by e.g. Funda [Eng] (but do not touch the entry "v" after "Category") > ok
   Repeat the procedure: e.g. Funda [Eng], Esra [Eng], Zehra [Eng] > ok

b. *Format Translation tier ((Arial Unicode MS, 10 points))*
   > activate a speaker's translation tier by mouse click into the speaker's translation tier label > Menue > Format > Format tier (or command + F) > Edit tier format
   > Font: Arial, Font style: Plain, Font size: 10 [you may also use properties *Text colour, Background colour, Alignment*] > ok

3) Morphological translation tier, Type: A(nnotation)

a. *Create morphological translation tier*
   > Menue > Tier > Add tier (command + A; or click icon) > Speaker: [no speaker], Type: A(nnotation) > ok (a new tier [v] appears in partitur) > Activate new tier [v] by mouse click into the name field > Menue > Tier > Edit tier properties > speaker: [no speaker] > Type: A(nnotation) > Category: v > Display: > cancel tickmark in square "Auto" and erase v after "Display"; re-

place it by e.g. Funda [TL] (but do not touch the entry “v” after “Category”) > ok
Repeat the procedure: e.g. Funda [TL], Esra [TL], Zehra [TL] > ok

b. *Format morphological translation tier*
> activate a speaker’s translation tier by mouse click into the speaker’s name translation field > Menue > Format > Format tier (or command + F) > Edit tier format > Font: Arial Unicode MS, 8 points > ok

((if necessary, you change allocated properties, later on, using the same procedure))

4) Tiers for comments, nonverbal communication (nvc), Type: D(escription)

a. *speakerspecified tier [nv]* for comments, nvc

1. Create speakerspecified nonverbal tier [nv]
> Menue > Tier > Add tier (command + A; or click icon) > Speaker: [no speaker], Type: D(escription) > ok (a new tier [v] appears in partitur) > activate tier label of new tier [v] and use “Move tier upwards” (via > Menue > Tier > Move tier upwards to the resepctive speaker’s tier label; or click icon with arrow pointing upwards or Command + upward cursor)

> Activate new tier [v] by mouse click into the name field > Menue > Tier > Tier properties > Edit tier properties > speaker: [no speaker] > Type: D(escription) > Category: v > Display: > cancel tickmark in square “Auto” and erase v after “Display”; replace it by e.g. Funda [nv] (but do not touch the entry “v” after “Category”) > ok
Repeat the procedure for other speakers, here e.g. for Esra [nv], Zehra [nv] > ok

2. *Format speakerspecified tier [nv*]
for comments and nonverbal communication which accompanies verbal communication of a specific speaker > activate a speaker’s *speakerspecified tier [nv*] by mouse click into the speaker’s tier label, e.g. Funda [nv] > Menue > Format > Format tier (or command + F) > Edit tier format > Font: Arial Unicode MS, Plain, 10 points > ok

b. *speakerunspecified (nonverbal) tier [nn], Type: D(escription)*
comments of those acoustic phenomena which can not to be attributed to a speaker, e.g. a being slammed, car noise, unidentifiable voices in the background etc.

1. Create speakerindependent tier [nn]
> Menue > Tier > Add tier (command + A; or click icon) > Speaker: [no speaker], Type: D(escription) > ok (a new tier [v] appears in partitur at lowest tier label level
> Activate new tier [v] by mouse click into the tier label > Menue > Tier > Tier properties > Edit tier properties > speaker: [no speaker] > Type: D(escription) > Category: v > Display: > cancel tickmark in square “Auto”

and erase v after “Display”; replace it by [nn] (but do not touch the entry “v” after “Category”) > ok > [nn| appears in partitur

2. *Format speakerindependent tier [nn]*
> activate *speakerindependent tier [nn]* by clicking into the respective tier label [nn] > Menue > Format > Format tier (or command + F) > Edit tier format > Font: Arial Unicode MS, Italics, 8 points > ok

VIII. *Formatting tier labels*

(domains/fields separately denominating each tier on the **left separate side-board** of the tiers of the score)
Use a unified format for all tier labels, e.g. Font: Arial, Font style: Plain, Font size: 10 pts

Formatting tier labels, proceed top -> down as follows:
> activate all tier labels of a partitur at once by clicking into them (standard procedure) > Menue > Format > Format tier labels (or click icon with hammer and yellow strip) > Edit tier format > e.g. Font: Arial, Font style: Plain, Font size: 10, Alignment: Left > ok

A note concerning default tier font

It happens that the previously installed setting, the preferences, for font type *Arial* in tiers may be changed, e.g. in case of ending up the program by crash. Then re-install preferences:
Menue > Edit > rider "Fonts" > Default tier font: Change > Set the default tier > Arial Unicode MS > ok > Preferences ok
Of course, you also can apply this procedure in order to install another font as default tier font.

IX. *Intonation tiers*

a. emphasis by (sublinearly) underlining the respective parts of expression:

Menue > Tier > Add tier... ((instead command+A)) > new tier
> ((allocate the additional tier to one of the speakers:)) Speaker: e.g. SPK0[ULF] > Type: A(nnotation) **> ok**

> ((determine properties of emphasis tier:)) activate the respective speaker and tier of the speaker in the open file > tier > tier properties > compare entries under Speaker, Type and Category > cancel entries under Display > cancel tickmark under Auto > ok

Move emphasis (empty) tier (sublinearly) under the appropriate speaker tier > activate the emphasis (empty) tier by a click in the speaker field > Menue > Format > Format tier ... (command+F) > Edit tier format > Font Arial: Arial > Font style: Plain > Font size: 9 > Background > click > colour table: click in the black square left/below > ok

Height of (intonation) tier
> Height calcul... > Fixed 1 Pixels > ok
((save data of emphasis tier as an own file and allocate them to the emphasis tier:)) > opened „Edit tier format“ > activate the emphasis (empty) tier by a click in the speaker field > Menue > Format > save format table as > name of the subfile into the window > allocate folder > ok

X. *Transcription symbols, utterance end marking signs for segmentation*

Utterance end symbol: *Transcription unit* is – according to HIAT transcription system – the ‘utterance‘. An utterance is to be marked by an utterance end symbol + space. By means of utterance end symbols, the transcript is segmented according to utterances, the transcribed and marked utterance is a ‘segment’. Single words may be utterances, too.
The following symbols normally are used (s. table below “Table of symbols used for utterance end marking (segmentation)

name of the symbol	orthographic sign	explanation
utterance end symbol	. ! ? … ˙	these symbols mark the end of utterance; the marking is obligatory in computer based transcripts; … an utterance is ended by being broken off; ˙ ends an interjection (incitement)
space symbol		follows obligatorily every utterance end symbol
quotations marks	“”	quotations marks are used before and after a quote, but not as utterance end symbols. In case that an utterance ends when a quote ends, one of the utterance end symbols above are used.
double parentheses (open, closed)	(( ))	count as non phonological segment and are treated as utterances
miscellaneous punctuation	‘ : ; , - _ ◡ — /	– marks an insertion within an utterance; / marks repairs, exactly speaking, the shifts from reparandum to the reparans within in an utterance; (French liaison) indicates the immediate connection(spoken without any pause) of a subsequent utterance to the following utterance

pause symbols	•	very short but audible
	• •	short
	• • •	half a second
	((1 s)), ((2 s))	mark 1 second, 2 second pauses and further

Table of symbols used for utterance end marking (segmentation)

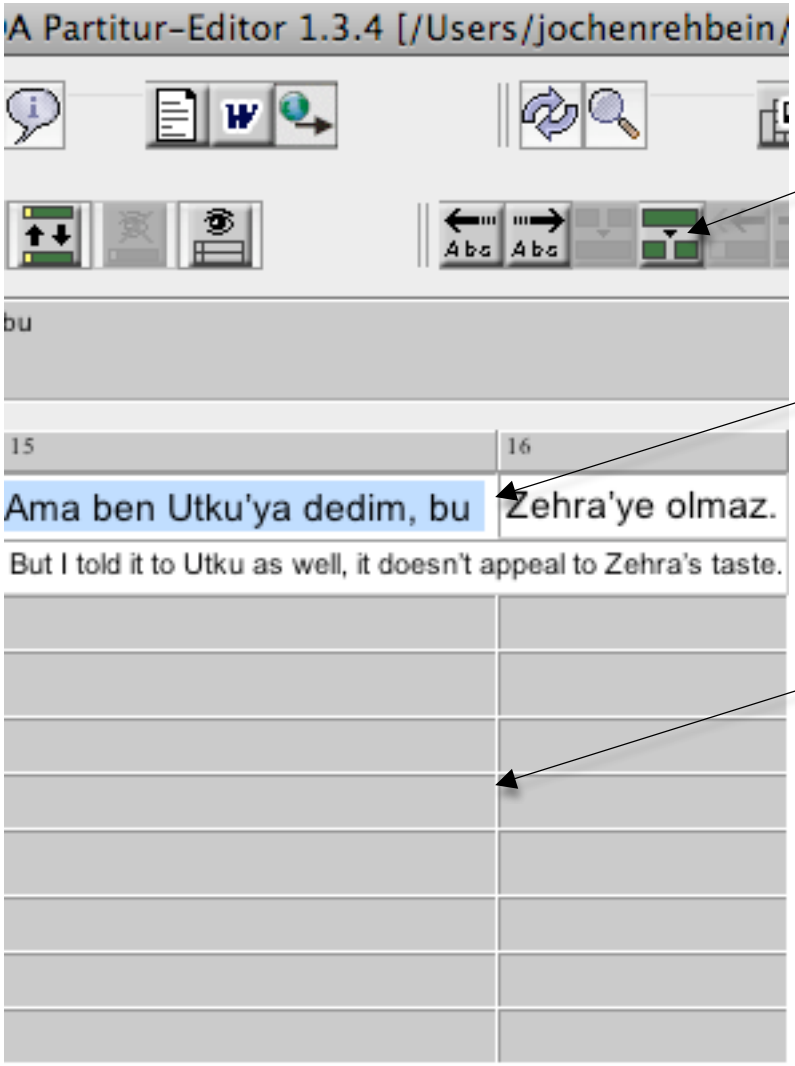
XI. *Keybord for Turkish and HIAT symbols for the verbal tiers*

> Menue > View > Show panels >((select:)) HIAT + Turkish supplement
(in the example, chose this keyboard for the tiers for Funda [v], Ezra [v] and Zehra [v])
Russian: select Cyrillic Alphabet
other languages, e.g. Georgian, Greek etc., are pre-constructed
((If necessary, you can change keyboard selection according to the respective signs needed under work))

XII. *Transcribing*

((For matters of describing the simultaneity of speech, we distinguish in the following between “tier of departure” where cursor is positioned and “target tier” in which the operation is effected))

Indicate overlapping simultaneity of parts of tiers (use split and merge)
Put the cursor in a position of a certain verbal tier (= “tier of departure”) > Menue > Event > Split (or click split icon or command + 2) > a new vertical time line is created showing a new number > the simultaneity is valid for the whole score area of all tiers > Start transcription in the tier referred to (= “target tier”) of another speaker; start at the right side of the new time line

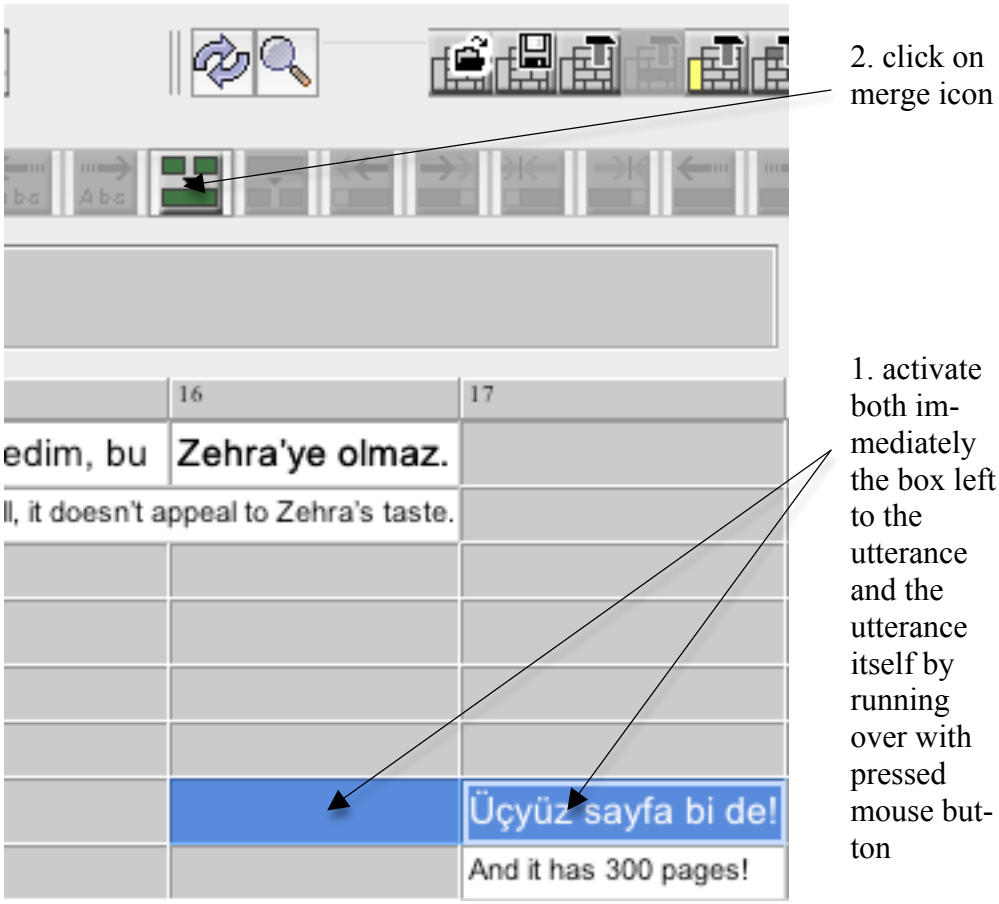


2. click on split icon

1. set the cursor into a certain tier („tier of departure)

3. A new vertical line is created with a new number in the time line (here: 16)

If the utterance in the target tier is transcribed already, then use merge
> Activate both the space left to the utterance (which emerged because of the new time line) and the already transcribed utterance
> Menue > Event > Merge (or click *merge icon* or *command* + *1*) > Transcribed utterance is transferred to the left
> Repeat operation with corresponding translation tier, transliteration tier or other respective tiers which are connected with the same utterance of the the same speaker
(the following two screen shots illustrate process and result of the merge operation)

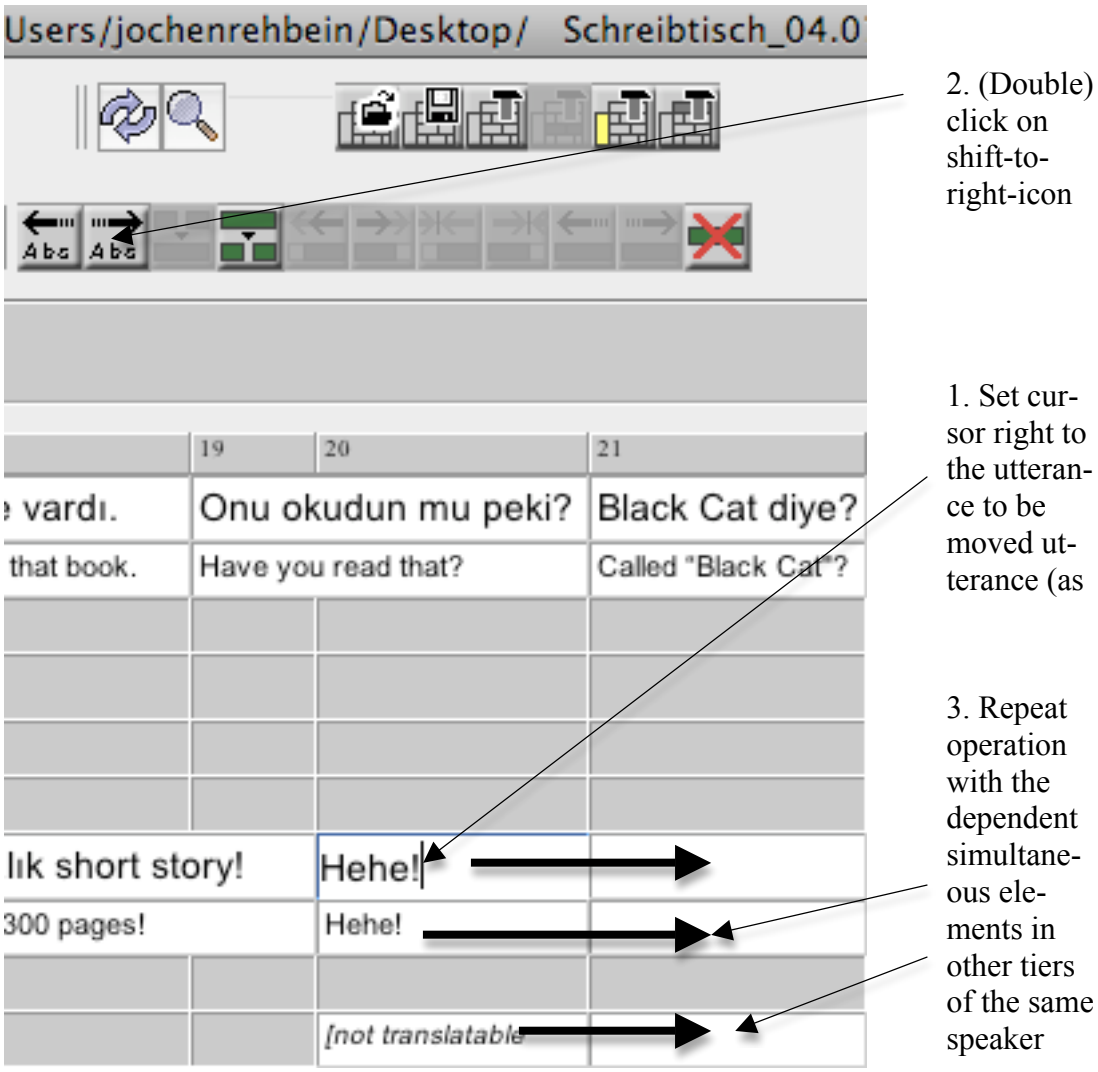


the transcribed utterance (as whole) is transferred to the left margin of the (new) box created by split operation; repeat the same operation with the corresponding translation tier

Movement of utterance(elements) to the right

Move characters, words (utterance elements) or the whole utterance over the next timeline to the right (use Shift characters to the right" or shift-to-right icon or command+shift+R)

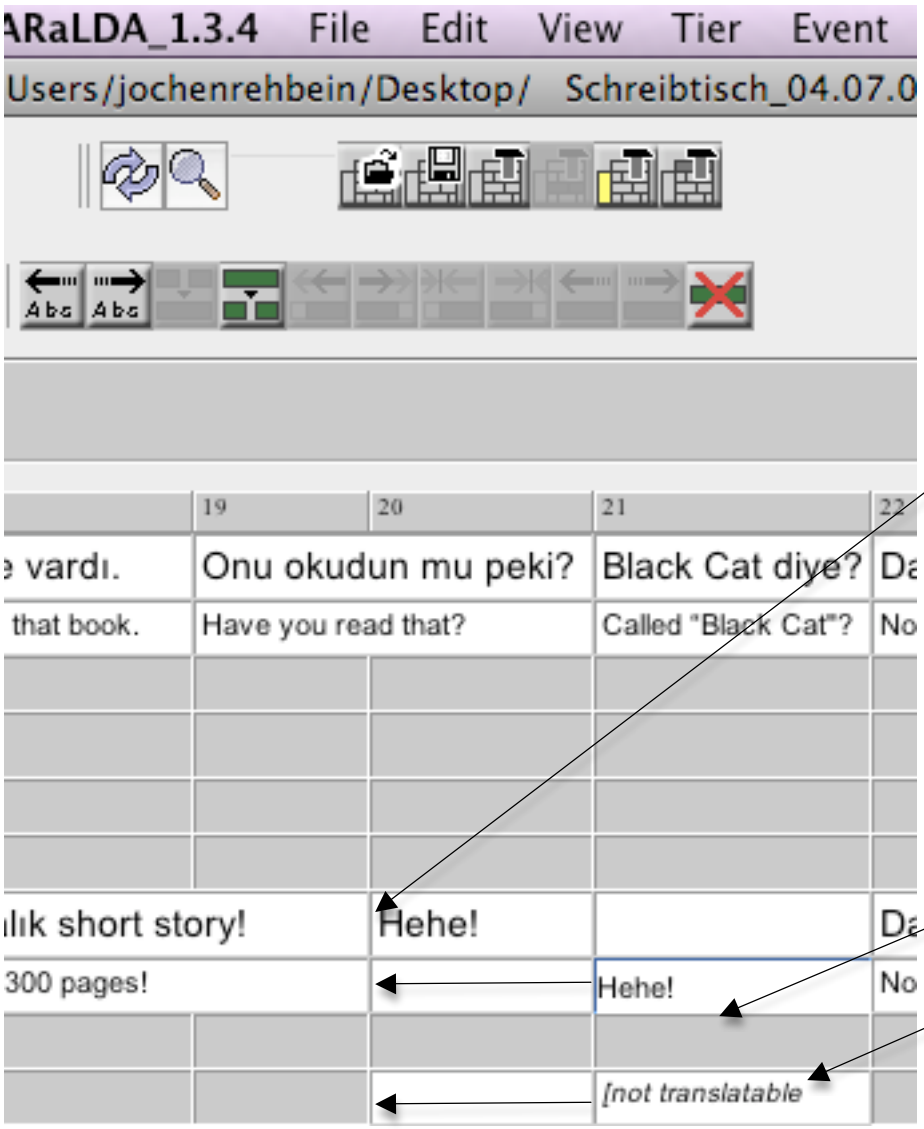
> Set cursor right to the utterance elements to be moved > Menue > Event > Event properties > Shift characters to right > utterance elements are moved over the next time line on the right



Utterance elements have been moved over the next time line to the right

Movement of utterance(elements) to the left
Push characters, words (utterance elements) or the whole utterance over the next timeline ***to the left*** *(*use *Shift characters to the left"* or *command+shift+R)*

> Set cursor right to the utterance elements to be pushed > Menue > Event > Event properties > Shift characters to left > utterance elements are moved over the next time line to the left



2. Menue > Event > Event properties > Shift characters to the left

3. Utterance elements are moved to the left

1. Set cursor right to the utterance to be moved

Moving the whole event (utterance) Shrink
((further functions are to be described here))

Use "remove unused timeline items" after having ended the transcript

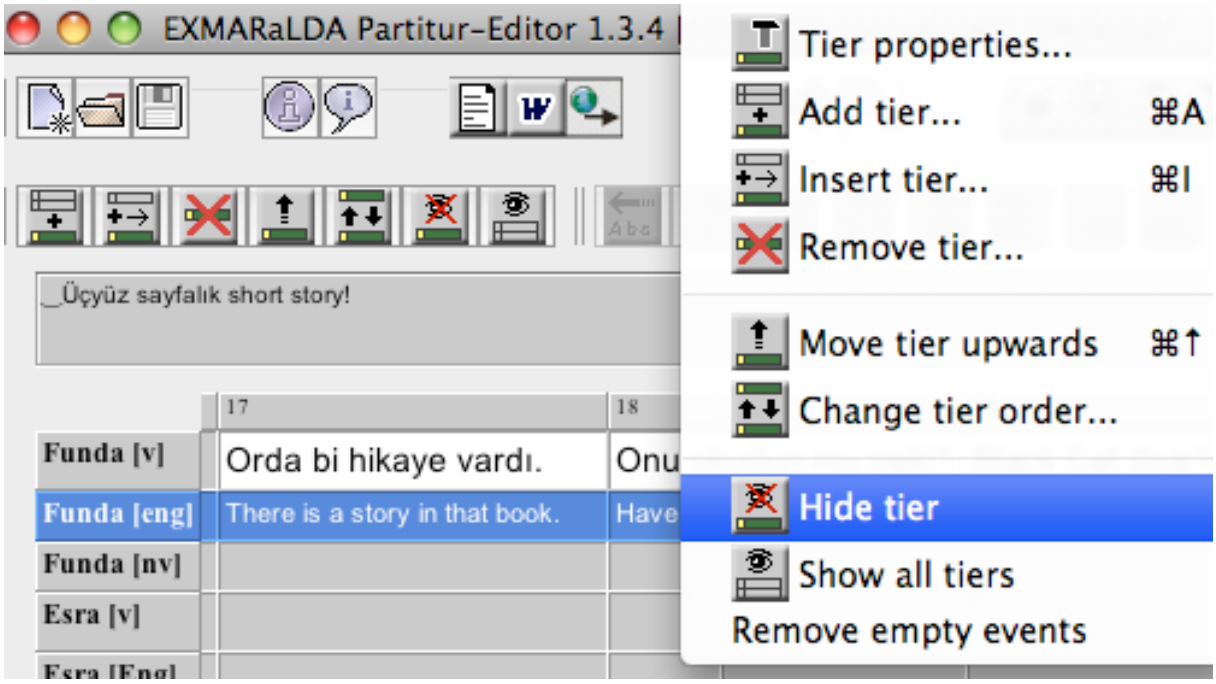
A note on the timelines of original verbal utterance and its (sublinear) translation
An original verbal utterance and its (sublinear) translation must be put into the corresponding timeline field, i.e. it must be limited by the same timelines at the beginning and at the end. You can use the merge function to put the translation into the corresponding timeline position.

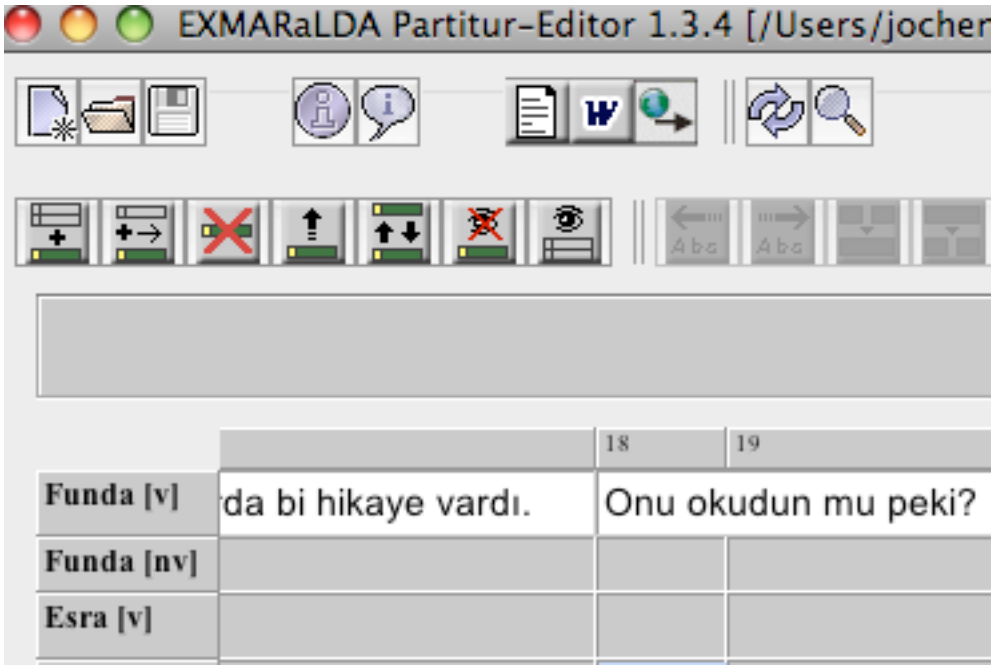
A note on the apperance of spaces or distances between words or characters caused by vertical time line
The precise indication of simultaneity (overlapping) of utterances by timelines should not expand the normal orthographical distance between words and/or characters. Therefore, use the possibility of moving the timeline by mouseclick and drawing it into the respective position appropriate to close the spaces and distances. (This operation is to be made for reasons of better visuability of overlapping linguistic forms.)

XIII. Suppress filled in tiers for print or presentation output

If you want to suppress a tier you have filled in with a translation, a morphological transliteration, a description of non verbal communication etc. – mostly a tier of an A-type or D-type – you proceed as follows:

> activate the respective tier by clicking into label name in the left side board of the score > Menue > Tier > Hide Tier > the activated tier disappears (but is not erased)





If you want to visualize suppressed tiers:
>Menue > Tier > Show all tiers > Disappeared tiers reappear

This function is recommendable, if you present the transcript to (e.g. linguistically) different audiences / reader groups.

For means of output in an RTF-format (export format into a document), use the following path:
Menue > Edit > Selection > w Selection to RTF > the whole score area (partitur) will be issued as RTF-document

XIV. *Visualizing (print, transfer and integration into text documents, e.g. in a word format)*

0. Preparing for visualizing
Menue > File > Partitur Parameters > choose parameters

1. Print Partitur
a) > Menue > File> Print (command + P) > (preview >) print
OR
b) > save as pdf-document or postscript file
> cut out of pdf-document or postscript file Partitur / score areas wanted and integrate them into word file

2. Partitur as RTF-Document [good print quality]
> Menue > File> Visualize > rtf-partitur (or click icon ***w*** in upper toolbar or use *command+R*) > save as *xxx* on hard disc > draw saved *xxx.rtf*-file on program-icon of word ((Note: Do not open saved *xxx.rtf*-file directly via the open-function of an opened word document!))> *xxx.rtf* file opens as *word.rtf* file
> embed Partitur (score) areas into your selected word document
OR
> print out *word xxx.rtf* file
OR
> save *word xxx.rtf* file *as xxx.pdf* file

Note
Do not use function “calculate page breaks” in Partitur Parameters, rider “RTF”, Macintosh system 10.5. upwards

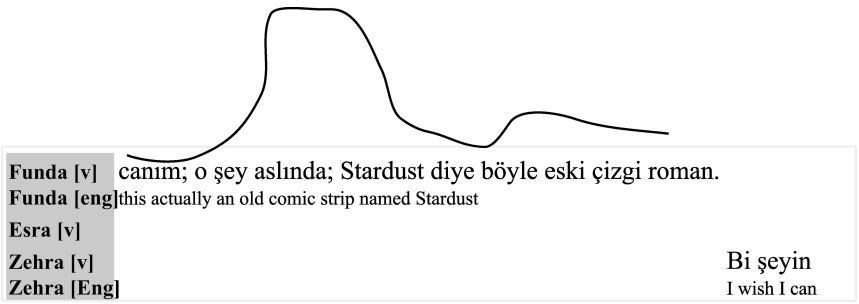
3. Partitur(s) as SVG-document(s)
(a) > Menue > File> Visualize > SVG-partitur > file name for superordinated HTML-File -> Sichern unter: xxxx.html -> save name of directory; subdirectory for SVG.html files (name.number. svg) save on your hard disc

> “You can open the file “SVG-Ausgabe.html” by means of any browser (e.g. Internet Explorer, Mozilla) which disposes of an appropriate SVG-Plugin (e.g. from Adobe or from Corel). Moreover, you also can open the individual SVG-files by means of an appropriate

software (e.g. Adobe Illustrator) and within another graphic format (e.g. WMF), save and embed them into word documents." (Schmidt 2006, p. IV-13)

(b) Print of SVG.html files via Inkscape

> Download Inkscapefor your system: http://www.inkscape.org/
> open Inkscape
> press down ctrl, double click (successively) on SVG files on your hard disc OR open SVG files via inkscape
> in inkscape via Datei > export Bitmap (ctrl + shift +E) > Bitmap-Größe einstellen: width at least Pixel 1000,00 dpi > export -> a .png file will be created (edited in the same folder as the SVG file)
> save .png file in Word Office as .rtf-file
> embed .rtf-files into your Word documents
> repeat procedure for all SVG files you wish to embed into your document



OR
> open .png file in Adobe Acrobat (if possible version 7 or 8) as .pdf-file (of an excellent print quality) [when embedded into word documents high quality of .pdf file remains]

> repeat procedure for all SVG files you wish to embed into your document

> In Inskape, you can draw prosodic curves above verbal utterances of the Partitur / score > and can save and print them in this format; s. example above.

References

Çilbir, Gamze (2008) Talking among friends (Gamze_TAF_150508). Audiotape recording 150508. Ankara: Middle East Technical University

Ehlich, K. & Rehbein, J. (1976) Halbinterpretative Arbeitstranskriptionen (HIAT). In: Linguistische Berichte 45/1976, 21-41

Ehlich, K.; Mackenzie, L.; Rehbein, J.; Thielmann. W. and Jan D. ten Thije (2006) A German-English-Dutch Glossary for Functional Pragmatics (mimeo)

Rehbein, J. (1995) Segmentieren. Zur Analyse von Einheiten der gesprochenen Sprache auf der Basis computergestützten Transkribierens. Universität Hamburg: Memo 64 des Verbundvorhabens Verbmobil. In: Rehbein, Jochen & Shinichi Kameyama, Bausteine diskursanalytischen Wissens. Berlin, New York: Mouton de Gruyter (in preparation 2009)

Rehbein, J.; Schmidt, T.; Meyer, B.; Watzke, F. & Herkenrath, A. (2004) Handbuch für das computergestützte Transkribieren nach HIAT. In: Arbeiten zur Mehrsprachigkeit, Folge B 56. University of Hamburg: Collaborative Research Institute (SFB 538) Multilingualism

Schmidt, T. (2006) EXMARaLDA. Partitur-Editor Handbuch. Version 1.3.2. University of Hamburg: Collaborative Research Institute (SFB 538) Multilingualism. http://www1.uni-hamburg.de/exmaralda/files/handbuch-aktuell.pdf