

# **EXMARaLDA – Creating, analysing and sharing spoken language corpora for pragmatic research**

Thomas Schmidt  
SFB 538 ‚Multilingualism‘  
University of Hamburg  
[thomas.schmidt@uni-hamburg.de](mailto:thomas.schmidt@uni-hamburg.de)

# Background

- Special Research Centre on Multilingualism, University of Hamburg
- 18 projects studying
  - multilingual language acquisition
  - multilingual communication
  - diachronic aspects of multilingualism
- Empirical approach using spoken and written language corpora
- Our project: developing methods for computer-assisted creation and analysis of multilingual language corpora

# EXMARaLDA

Extensible Markup Language for Discourse Annotation

Main aims:

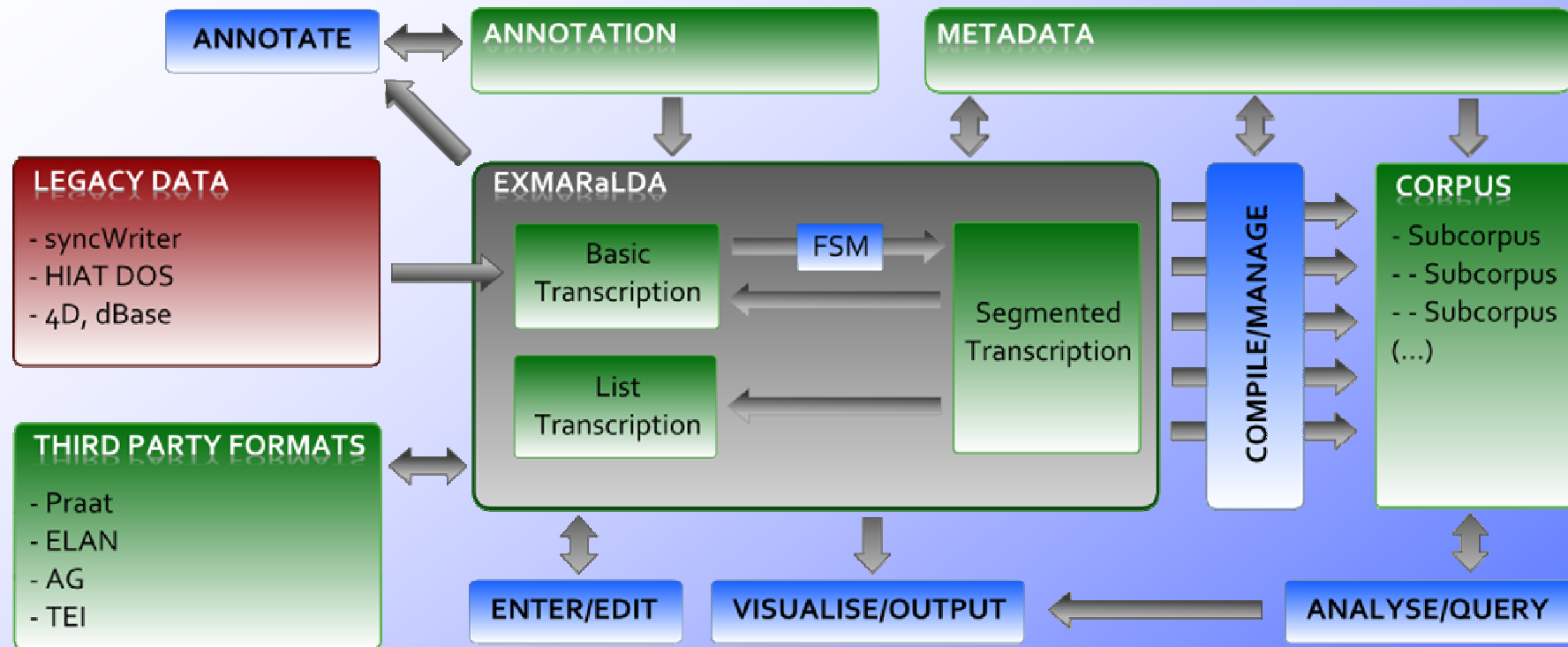
- to facilitate the **exchange** of spoken language corpora between researchers and between technological environments,
- to optimally exploit the **multimedia and hypertext** capabilities of modern computers in the work with video or audio data and their transcriptions,
- to pave the way for **long term archiving** and reuse of costly and valuable language resources.

# EXMARaLDA

Used in:

- Language acquisition studies
- Conversation / Discourse Analysis, Functional Pragmatics
  - Complexity of spontaneous spoken language:  
**multi-party, multi-level, multi-modal**
  - Importance of **context**: interactional and situational context, ethnographic meta-data
  - Explorative, **corpus-driven** analyses

# EXMARaLDA system architecture



- Data centric
- Using a time-based data model (cf. Annotation Graphs, Bird/Lieberman)
- Using open standards: XML and Unicode
- Using cross-platform technology: Java
- Tools for transcription, corpus management, corpus query and analysis

# Transcription: Partitur-Editor

EXMARaLDA Partitur-Editor 1.3.4 [S:\TP-Z2\DATEN\BEISPIELE\LYON\Pear Story\BasicTransPear.xml]

File Edit View Tier Event Timeline Format Segmentation Help

Okay.

	16 [17.0]	17	18 [18.9]	19 [19.4]	20 [19.9]	21 [20.4]	22 [21.0]	23 [22.4]
					louder			
X [v]	And he puts the pears into an		a	pron.	Okay?	((takes breath, 0,5s))	Aand so then	he gathers a bunch.
X [mv1]	<i>bFA simultaneously, up and down in front of BR</i>			<i>bHA rest in front of BR</i>		<i>rHA rubs NO, lHA down on rKN</i>		<i>bFA move simultaneously in front of BR</i>
Y [v]			Okay.					

Done.

Keyboard

International Phonetic Alphabet

Open...

e	o	ɒ	b	ɔ	ç	q	ɔ	e
a	ɔ	ç	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ
g	y	x	u	h	ɔ	l	l	l
t	f	l	ɔ	w	u	m	n	ɔ
N	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ
t	r	ɔ	r	ɔ	ɔ	ɔ	ɔ	ɔ
f	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ
Y	z	z	z	z	z	z	z	z
o	B	o	ɔ	H	J	x	L	ɔ
?	ç	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ

Audio/Video panel

.\DATEN\BEISPIELE\LYON\Pear Story\pear.mov

Start

Position

Stop

0.0 19.5 19.5 43.8 43.8 sec

Playback stopped

# Transcription: Partitur-Editor

The screenshot displays the EXMARaLDA Partitur-Editor 1.3.4 interface. The main window shows a musical score notation with a table of transcription details. The table is highlighted with an orange border. Below the main window, there are two smaller panels: a Keyboard panel and an Audio/Video panel.

	16 [17.0]	17 [18.9]	18 [18.9]	19 [19.4]	20 [19.9]	21 [20.4]	22 [21.0]	23 [22.4]
					louder			
X [v]	And he puts the pears into an a pron.				Okay? ((takes breath, 0,5s))	Aand so then	he gathers a bunch.	
X [mv1]	<i>bFA simultaneously, up and down in front of BR</i>			<i>bHA rest in front of BR</i>		<i>rHA rubs NO, lHA down on rKN</i>		<i>bFA move simultaneously in front of BI</i>
Y [v]				Okay.				

**Keyboard panel:** Shows the International Phonetic Alphabet (IPA) keyboard with various phonetic symbols.

**Audio/Video panel:** Shows a video player with a scene of two people talking. The video is titled "DATEN\BEISPIELE\LYON\Pear\_Story\pear.mov". The playback is stopped at 0:00. The video duration is 43.8 sec.

# Transcription: Partitur-Editor

EXMARaLDA Partitur-Editor 1.3.4 [S:\VTP-ZZ\DATEN\BEISPIELE\LYON\Pear Story\BasicTransPear.xml]

File Edit View Ijer Event Timeline Format Segmentation Help

Okay.

16 [17.0] 19 [19.9] 20 [20.0] 23 [22.4]

**Audio and video support**

louder

X [v] And he puts the pears into an a pron. Okay? ((takes breath, 0,5s)) Aand so then he gathers a bunch.

X [mv1] bFA simultaneously, up and down in front of BR bHA rest in front of BR rHA rubs NO, lHA down on rKN bFA move simultaneously in front of BI

Y [v] Okay.

Done.

Keyboard

International Phonetic Alphabet

Grid of IPA symbols for transcription.

Audio/Video panel

.DATEN\BEISPIELE\LYON\Pear Story\pear.mov

Start: 0.0, 19.5, 19.5, 43.8, 43.8 sec

Position: 19.5

Stop: 43.8

0.0 19.5 19.5 43.8 43.8 sec

Playback stopped

Video window showing a man and a woman sitting and talking.

# Transcription: Output

[1]

0▶ 1 2▶ 3 4 5 6 7 8

A [v] Ja. Würde erstmal ganz gerne hören • wie es ihm geht.

P [v] Ah, eu sinto-me

P [de] Ah, ich fühle mich gut.

D [v] *gedehnt* Ehm • • o seu/ o, o senhor doutor tá-lhe a perguntar, como é que se sente? *lauter*

D [de] *Ähm Ihr/ der der Herr Doktor fragt Sie, wie Sie sich fühlen?*

[nn] ((Hintergrundgeräusche))

[2]

... 9 10 11 12.13 14▶ 15 16 17 18▶

A [v] Ihm geht's gut.

P [v] bem. Sim. • Sinto-me a calhar doente ((stottert, 10s)) mas c/c/ como eu sinto mal é que o/ ((1,5s)) é uma perna.

P [de] Ja. Ich fühle mich aber d/d/ da es mir nicht gut geht

D [v] Der sagt, ihm geht's gut.

[3]

18▶ 19 20 21 22 23 24 25▶

A [v] Hmhm'

P [v] É só das pernas. É só das pernas. Ja. • Und er zeigt auf seine

P [de] Es sind nur die Beine. Es sind nur die Beine.

D [v] Also, ähm (d)er fühlt sich krank, ähm gut aber andererseits krank wegen seiner Beine, die/ äh äh dass er nicht laufen kann.

- Different layouts,
- Different formats,
- Hypermedia

- ▶ 1 A Ja.
- 2 Würde erstmal ganz gerne hören • wie es ihm geht.
- ▶ 3 D Ehm • • o seu/ o, o senhor doutor tá-lhe a perguntar, como é que se sente? *ÄhmIhr/ der der Herr Doktor fragt Sie, wie Sie sich fühlen?*
- 4 P Ah, eu sinto-me bem. *Ah, ich fühle mich gut.*
- 5 D Der sagt, ihm geht's gut.
- 6 A Ihm geht's gut.
- 7 P Sim. *Ja.*
- 8 • Sinto-me a calhar doente ((stottert, 10s)) mas c/c/ como eu sinto mal é que o/ ((1,5s)) é uma perna. *Ich fühle mich aber d/d/ da es mir nicht gut geht*
- ▶ 9 D Also, ähm (d)er fühlt sich krank, ähm gut aber andererseits krank wegen seiner Beine, die/ äh äh dass er nicht laufen kann.
- 10 P É só das pernas. *Es sind nur die Beine.*
- 11 P É só das pernas. *Es sind nur die Beine.*
- 12 A Hmhm'
- ▶ 13 A Ja.

# Corpus (meta-data) management: CoMa

The screenshot displays the CoMa software interface for managing a corpus. The window title is "Coma2 | e5kopie2.xml | Corpus: E5 Korpus". The interface is divided into several panels:

- Filter:** Shows filters for "SKOBI" and "Uçuk", each with "Invert", "Add", and "Active" checkboxes. A button "Alle Filter löschen" is next to "28 Kommunikationen".
- Kommunikation:** A table listing communications with columns "S", "Name", and "Familie (Pseudo)".
- Eigenschaften:** A detailed view for "Communication 0848" with fields for "Id", "Description", "Thema Inhalt", "Projekt", "KassettenNr", "FileMakerID", "Familie (Pseudo)", "Diskursmodus", "DOrt räuml.", "Bearbeitungsstand", "Location", "Language(s)", and "Language".
- Personen:** A list of "zugeordnete Personen" with columns "S", "Sigle", and "Var".
- Aktion:** A panel with buttons for "Kommunikation hinzufügen", "Kommunikation löschen", "Alle angezeigten in den Korb", and "Segmente zählen".

S	Name	Familie (Pseudo)
<input type="checkbox"/>	0848	Uçuk
<input type="checkbox"/>	0723	Uçuk
<input type="checkbox"/>	0992	Uçuk
<input type="checkbox"/>	0931	Uçuk
<input type="checkbox"/>	0993	Uçuk
<input type="checkbox"/>	0661	Uçuk
<input type="checkbox"/>	0746	Uçuk
<input type="checkbox"/>	0678	Uçuk
<input type="checkbox"/>	0932	Uçuk
<input type="checkbox"/>	0990	Uçuk
<input type="checkbox"/>	0714	Uçuk
<input type="checkbox"/>	0755	Uçuk
<input type="checkbox"/>	1134	Uçuk
<input type="checkbox"/>	0991	Uçuk
<input type="checkbox"/>	0938	Uçuk
<input type="checkbox"/>	0722	Uçuk
<input type="checkbox"/>	0743	Uçuk
<input type="checkbox"/>	1169	Uçuk
<input type="checkbox"/>	1272	Uçuk
<input type="checkbox"/>	1296	Uçuk
<input type="checkbox"/>	1294	Uçuk
<input type="checkbox"/>	1455	Uçuk
<input type="checkbox"/>	1451	Uçuk
<input type="checkbox"/>	1423	Uçuk
<input type="checkbox"/>	1271	Uçuk
<input type="checkbox"/>	1294	Uçuk
<input type="checkbox"/>	1296	Uçuk
<input type="checkbox"/>	1447	Uçuk

S	Sigle	Var
<input type="checkbox"/>	Bir	Birgül
<input type="checkbox"/>	Ken	Kenan
<input type="checkbox"/>	Ann	Anna
<input type="checkbox"/>	Üns	Ünsal
<input type="checkbox"/>	Fer	Ferda
<input type="checkbox"/>	Cev	Cevat
<input type="checkbox"/>	Nes	Nesli
<input type="checkbox"/>	Mut	Mutter
<input type="checkbox"/>	Hai	Halime
<input type="checkbox"/>	Sam	Samet
<input type="checkbox"/>	Oyc	Oycan
<input type="checkbox"/>	Kam	Kamil
<input type="checkbox"/>	Mus	Musa
<input type="checkbox"/>	Vat	Vater
<input type="checkbox"/>	Mut	Mutter
<input type="checkbox"/>	Tug	Tuğba
<input type="checkbox"/>	Mek	Melike
<input type="checkbox"/>	Tan	Taner
<input type="checkbox"/>	Sef	Sefer
<input type="checkbox"/>	Vat	Vater
<input type="checkbox"/>	Mut	Mutter
<input type="checkbox"/>	Tah	Tahir
<input type="checkbox"/>	Bes	Beste
<input type="checkbox"/>	Vat	Alican
<input type="checkbox"/>	Mut	Şükriye
<input type="checkbox"/>	Sed	Seda
<input type="checkbox"/>	Şey	Şeyda
<input type="checkbox"/>	Kub	Kubat
<input checked="" type="checkbox"/>	Bin	Binnaz
<input checked="" type="checkbox"/>	Bur	Burak
<input type="checkbox"/>	Ben	Bengi
<input type="checkbox"/>	Han	Handan

# Corpus (meta-data) management: CoMa

## Communication and speaker list

The screenshot displays the CoMa software interface for managing a corpus. The main window shows a list of communications and a detailed view of a selected communication (ID: 0848). Two tables are highlighted with orange borders:

**Communication List (Left):**

S	Name	Familie (Pseudo)
<input type="checkbox"/>	0848	Uçuk
<input type="checkbox"/>	0723	Uçuk
<input type="checkbox"/>	0992	Uçuk
<input type="checkbox"/>	0931	Uçuk
<input type="checkbox"/>	0993	Uçuk
<input type="checkbox"/>	0661	Uçuk
<input type="checkbox"/>	0746	Uçuk
<input type="checkbox"/>	0678	Uçuk
<input type="checkbox"/>	0932	Uçuk
<input type="checkbox"/>	0990	Uçuk
<input type="checkbox"/>	0714	Uçuk
<input type="checkbox"/>	0755	Uçuk
<input type="checkbox"/>	1134	Uçuk
<input type="checkbox"/>	0991	Uçuk
<input type="checkbox"/>	0938	Uçuk
<input type="checkbox"/>	0722	Uçuk
<input type="checkbox"/>	0743	Uçuk
<input type="checkbox"/>	1169	Uçuk
<input type="checkbox"/>	1272	Uçuk
<input type="checkbox"/>	1296	Uçuk
<input type="checkbox"/>	1294	Uçuk
<input type="checkbox"/>	1455	Uçuk
<input type="checkbox"/>	1451	Uçuk
<input type="checkbox"/>	1423	Uçuk
<input type="checkbox"/>	1271	Uçuk
<input type="checkbox"/>	1294	Uçuk
<input type="checkbox"/>	1296	Uçuk
<input type="checkbox"/>	1447	Uçuk

**Speaker List (Right):**

S	Sigle	Var
<input type="checkbox"/>	Bir	Birgül
<input type="checkbox"/>	Ken	Kenan
<input type="checkbox"/>	Ann	Anna
<input type="checkbox"/>	Üns	Ünsal
<input type="checkbox"/>	Fer	Ferda
<input type="checkbox"/>	Cev	Cevat
<input type="checkbox"/>	Nes	Nesli
<input type="checkbox"/>	Mut	Mutter
<input type="checkbox"/>	Hai	Halime
<input type="checkbox"/>	Sam	Samet
<input type="checkbox"/>	Oyc	Oycan
<input type="checkbox"/>	Kam	Kamil
<input type="checkbox"/>	Mus	Musa
<input type="checkbox"/>	Vat	Vater
<input type="checkbox"/>	Mut	Mutter
<input type="checkbox"/>	Tug	Tuğba
<input type="checkbox"/>	Mek	Melike
<input type="checkbox"/>	Tan	Taner
<input type="checkbox"/>	Sef	Sefer
<input type="checkbox"/>	Vat	Vater
<input type="checkbox"/>	Mut	Mutter
<input type="checkbox"/>	Tah	Tahir
<input type="checkbox"/>	Bes	Beste
<input type="checkbox"/>	Vat	Alican
<input type="checkbox"/>	Mut	Şükriye
<input type="checkbox"/>	Sed	Seda
<input type="checkbox"/>	Şey	Şeyda
<input type="checkbox"/>	Kub	Kubat
<input checked="" type="checkbox"/>	Bin	Binnaz
<input checked="" type="checkbox"/>	Bur	Burak
<input type="checkbox"/>	Ben	Bengi
<input type="checkbox"/>	Han	Handan

The background interface shows the details for Communication 0848, including its description, project information (SKOBI), cassette number (0848), file maker ID (1196), family (Uçuk), discourse mode (EFE04 tk - Maulwurf), location (Hamburg), and language (TRK TR).

# Corpus (meta-data) management: CoMa

## Meta data

The screenshot displays the CoMa software interface for managing a corpus. The main window is titled 'Coma2 | e5kopie2.xml | Corpus: E5 Korpus'. The interface is divided into several panels:

- Filter:** Shows filters for 'SKOBI' and 'Uçuk', both with 'Active' checked. A button 'Alle Filter löschen' is visible.
- Kommunikation:** A table listing communications with columns 'S', 'Name', and 'Familie (Pseudo)'. The table contains 28 rows, all with 'Uçuk' as the family name.
- Eigenschaften (Properties):** A detailed view for 'Communication 0848' (ID: IDCommunication1197). It includes:
  - Description (Communication):** A text description: 'Das Gespräch findet im Kinderzimmer statt. Izel nimmt zuerst Bruder dran. Er redet darüber, was er auf dem Bild sieht. Izel stellt ab und zu Fragen. Dann kommt Binnaz ins Zimmer rein. Sie erzählt lange über die Bildfiguren. Dann vergleicht Binnaz sie mit den Filmfiguren, die sie im Fernsehen gesehen hat. Somit endet die Aufnahme.'
  - Thema Inhalt:** A section for content-related metadata.
  - Projekt:** SKOBI
  - KassettenNr:** 0848
  - FileMakerID:** 1196
  - Familie (Pseudo):** Uçuk
  - Diskursmodus:** EFE04 tk - Maulwurf
  - DOrt räuml.:** Kinderzimmer
  - Bearbeitungsstand:** noch nichts passiert
  - Location:**
    - City:** Hamburg
    - Country:** D
    - PStart:** 24.01.2002 00:00:00
  - Description (Location):** (Empty)
  - Language(s):** (Empty)
  - Language:** TRK TR
  - Description (Language):** (Empty)
- Personen:** A list of people associated with the communication, including Bir, Ken, Ann, Uns, Fer, Cev, Nes, Müt, Hal, Sam, Oyc, Kam, Mus, Vat, Mut, Tug, Mek, Tan, SeF, Vat, Müt, Tah, Bes, Vat, Mut, Sed, Sey, Kub, Bin, Bur, Ben, Han, and others.

# Corpus query and analysis: EXAKT

EXAKT

File Edit

RECENTLY USED

Corpora

- Dolmetschen im Krankenhaus**
  - S:\TP-Z...ing\K2\_Corpus.xml
  - 111 transcriptions
  - 20710 segment chains
- K5\_Toerring**
  - S:\TP-Z...ing\K5\_Toerring.xml
  - 2 transcriptions
  - 3484 segment chains
- K5\_Oeresund**
  - S:\TP-Z...und\K5\_Oeresund.xml
  - 35 transcriptions
  - 7364 segment chains

Done.

Concordances

- eigentlich**
  - Dolmetschen im Krankenhaus
  - 79 tokens
  - 1 types
- det**
  - K5\_Toerring
  - 846 tokens
  - 1 types
- myck, mycket**
  - K5\_Oeresund
  - 208 tokens
  - 2 types
- ver-me, dizia-me, fui-me, Anda-me**
  - Dolmetschen im Krankenhaus
  - 152 tokens
  - 106 types

Dolmetschen im Krankenhaus (79 results) | K5\_Toerring (846 results) | K5\_Oeresund (208 results) | Dolmetschen im Krankenhaus (152 results)

RegEx (T) Search: \b[A-Za-z]+me\b

S	C...	Speaker	Left Context	Match	Right Context	Type	Diskursar...	Konstellat...
91	SGa		E eu	meti-me	no carro, feito maluco, vim soz		Aufklärung	Monolingual
92	INa		mentos. ((2s)) ((schnalzt)) Eu,	falta-me	um exame, ((1,5s)) falta-me um	exclamative	Aufklärung	Monolingual
92	INa		Eu, falta-me um exame, ((1,5s))	falta-me	um exame para-lhe/ para res	other	Aufklärung	Monolingual
92	DZé		liguei muito. ((1s)) Ainda me	deixei-me	andar. ...E depois aquilo eh pas		Aufklärung	Monolingual
83	IFre		E eu quem sou?	Conhece-	a mim?	declarative	Anamnese	Monolingual
94	DMar		((2s)) Olhe,	deu-me	((unverständlich, 1s)) e a dar-		Anamnese	Monolingual
94	DMar		u-me ((unverständlich, 1s)) e a	dar-me	assim tonturas pela cabeça. •I		Anamnese	Monolingual
94	DMar		eu médico lá de família ace/ eh	receitou-me	•• aconselhou-me para fazer te	interrogative	Anamnese	Monolingual
94	DMar		família ace/ eh receitou-me ••	aconselho-	para fazer terapia e fui fazer	interrogative	Anamnese	Monolingual
94	DMar		ara fazer terapia e fui fazer e	encontrei-	muito mal.		Anamnese	Monolingual
94	DMar		á bastante tempo, mas o segundo	deu-me	cá de uma maneira. ((holt hörba		Anamnese	Monolingual
94	DMar		/ ao girar assim com o pescoço,	estalar-me	aqui muito. Foi por isso que o	declarative	Anamnese	Monolingual
94	DMar		nturas não. Deitada não, mas ao	por-me	a pé, vou já cair.		Anamnese	Monolingual
101	ILu		((blättert, 2 s)) Ora	deixe-me	cá ver. ((blättert, 3,5 s)) Ist	interroa...	Befund	Monolingual
101	ILu		Eu percebi,	estou-me	a rir, mas percebi.	interrogative	Befund	Monolingual
101	Umb		que me tinha acabado, não é, e	veio-me	e depois nunca mais me veio. Ac	exclamative	Befund	Monolingual
18	Mar		Ah, eu	sinto-me	bem.	declarative	Aufnahme ...	Gedolmets...
18	Mar		Sim. •	Sinto-me	a calhar dnente ((stottert 10s	other	Aufnahme	Gedolmets

((blättert, 2 s)) Ora **deixe-me** cá ver. ((blättert, 3,5 s)) Isto hoje foi complicado. ((blättert, 2 s))

Type	interrogative
Diskursart[C]	Befund
Konstellationstyp[C]	Monolingual

Types: 106  
Tokens: 152  
Selected: 152  
Time: 32.14 s

Partitur

A [v] É? A senhora sabe algum? ((blättert, 2 s)) Ora **deixe-me** cá ver. ((blättert, 3,5 s)) Isto hoje foi complicado.

A [de] Ja? Wissen Sie welche? Also, lassen Sie mir sehen. Das heute war kompliziert.

P [v] É. Não é? É. Tá a ver?

P [de] Ja. Nicht wahr? Ja. Sehen Sie?

Partitur HTML

# Corpus query and analysis: EXAKT

The screenshot displays the EXAKT software interface. A search bar at the top center contains the regular expression `lb[A-Za-z]+me\b` and is highlighted with an orange box. Below the search bar, a table of search results is shown, with the row containing `deixe-me` highlighted in blue. The table columns include line number, speaker, text, search results, and linguistic annotations. A large white text overlay reads "Regular expression search".

S	C...	Speaker	Text	Search Results	Type	Context	Language
91		SGa	E eu	mett-me		no carro, feito maluco, vim soz	Aufklärung Monolingual
92		INa	mentos. ((2s)) ((schnalzt)) Eu,	falta-me	exclamative	um exame, ((1,5s)) falta-me um	Aufklärung Monolingual
92		INa	Eu, falta-me um exame, ((1,5s))	falta-me	other	um exame para-lhe. * para * res	Aufklärung Monolingual
92		DZê	*liguei muito. ((1s)) Ainda me	dexei-me		andar. _E depois aquilo eh pas	Aufklärung Monolingual
93		IFre	E eu quem sou?	Conhece-	declarative	a mim?	Anamnese Monolingual
94		DMar	((2s)) Olhe,	deu-me		((unverständlich, 1s)) e a dar-	Anamnese Monolingual
94		DMar	u-me ((unverständlich, 1s)) e a	dar-me		assim tonturas pela cabeça. *	Anamnese Monolingual
94		DMar	eu médico-lá de família acef/eh	receitou-me	interrogative	* * aconselhou-me para fazer te	Anamnese Monolingual
94		DMar	família acef/eh receitou-me *	aconselho	interrogative	para fazer terapia e fui fazer	Anamnese Monolingual
94		DMar	ara fazer terapia e fui fazer e	encontrei-		muito mal	Anamnese Monolingual
94		DMar	á bastante tempo, mas o segundo	deu-me		ca de uma maneira. ((hoit horba	Anamnese Monolingual
94		DMar	/ ao girar assim com o pescoço,	estalar-me	declarative	aqui muito. Foi por isso que o	Anamnese Monolingual
94		DMar	nturas não. Deritada não, mas ao	por-me		a pé, vou já cair.	Anamnese Monolingual
101		ILu	((blãtert, 2 s)) Ora	deixe-me	interrog...	cã ver. ((blãtert, 3,5 s)) Ist	Befund Monolingual
101		ILu	Eu percebi,	estou-me	interrogative	a rir, mas percebi.	Befund Monolingual
101		Umb	que me tinha acabado, não é, e	veio-me	exclamative	e depois nunca mais me veio. Ac	Befund Monolingual
18		Mar	Ah, eu	sinto-me	declarative	bem.	Aufnahme Gedolmets
18		Mar	Sim *	Sintu-me	other	a calhar dnente ((sintert 10s	Aufnahme Gedolmets

Regular expression search

# Corpus query and analysis: EXAKT

EXAKT

File Edit

RECENTLY USED

Corpora

Dolmetschen im Krankenhaus (79 results) KS\_Toerring (846 results) KS\_Oeresund (208 results) Dolmetschen im Krankenhaus (152 results)

Search: [bfa-Za-z] me

S	C...	Speaker	Left Context	Match	Right Context
91	SGa			E eu <b>meti-me</b>	no carro, feito maluco, vim soz
92	INa		mentos. ((2s)) ((schnalzt)) Eu,	<b>falta-me</b>	um exame, ((1,5s)) falta-me um
92	INa		Eu, falta-me um exame, ((1,5s))	<b>falta-me</b>	um exame para-lhe/ para res
92	DZé		liguei muito. ((1s)) Ainda me	<b>deixei-me</b>	andar. ...E depois aquilo eh pas
83	IFre		E eu quem sou?	<b>Conhece-me</b>	a mim?
94	DMar		((2s)) Olhe,	<b>deu-me</b>	((unverständlich, 1s)) e a dar-
94	DMar		u-me ((unverständlich, 1s)) e a	<b>dar-me</b>	assim tonturas pela cabeça. I
94	DMar		eu médico lá de família ace/ eh	<b>receitou-me</b>	•• aconselhou-me para fazer te
94	DMar		família ace/ eh receitou-me ••	<b>aconselho-me</b>	para fazer terapia e fui fazer
94	DMar		ara fazer terapia e fui fazer e	<b>encontrei-me</b>	muito mal.
94	DMar		á bastante tempo, mas o segundo	<b>deu-me</b>	cá de uma maneira. ((holt hörba
94	DMar		/ ao girar assim com o pescoço,	<b>estalar-me</b>	aqui muito. Foi por isso que o
94	DMar		nturas não. Deitada não, mas ao	<b>por-me</b>	a pé, vou já cair.
101	ILu		((blättert, 2 s)) Ora	<b>deixe-me</b>	cá ver. ((blättert, 3,5 s)) Ist
101	ILu		Eu percebi,	<b>estou-me</b>	a rir, mas percebi.
101	Umb		que me tinha acabado, não é, e	<b>veio-me</b>	e depois nunca mais me veio. Ac
18	Mar		Ah, eu	<b>sinto-me</b>	bem.
18	Mar		Sim.	<b>Sinto-me</b>	a calhar doente ((stottert, 10s

Concordances

eigentlich  
Dolmetschen im Krankenhaus  
79 tokens  
1 types

det  
KS\_Toerring  
846 tokens  
1 types

myck, mycket  
KS\_Oeresund  
208 tokens  
2 types

ver-me, dizia-me, fui-me, Anda-me  
Dolmetschen im Krankenhaus  
152 tokens

Partbur

A [v] É? A senhora sabe algum? ...e cá ver. ((blättert, 3,5 s)) Isto hoje foi complicado

A [de] Ja? Wissen Sie welche? Also, lassen Sie mich sehen. Das heute war kompliziert.

P [v] É. Não é? É. Tá a ver?

P [de] Ja. Nicht wahr? Ja. Sehen Sie?

Partbur HTML

KWIC concordance

# Corpus query and analysis: EXAKT

**EXAKT**

File Edit

RECENTLY USED

Corpora

Dolmetschen im Krankenhaus (79 results) KS\_Toerring (846 results) KS\_Oeresund (208 results) Dolmetschen im Krankenhaus (152 results)

RegEx (T) Search: b[A-Za-z]+me\b

S	C...	Speaker	Left-Context	Match	Right-Context	Type	Diskursar...	Konstellat...
91	SGa		E eu	mett-me	no carro, feito maluco, vim soz		Aufklärung	Monolingual
92	INa		mentos. ((2s)) ((schlätzt)) Eu,	falta-me	um exame, ((1,5s)) falta-me um	exclamative	Aufklärung	Monolingual
92	INa		Eu, falta-me um exame, ((1,5s))	falta-me	um exame para-lhe/* para *res	other	Aufklärung	Monolingual
92	DZé		*iguei muito. ((1s)) Ainda me ((	deixe-me	andar. _E depois aquilo eh pas		Aufklärung	Monolingual
83	IFre		E eu quem sou?	Conhece-	a mim?	declarative	Anamnese	Monolingual
94	DMar		((2s)) Olhe,	deu-me	((unverständlich, 1s)) e a dar-		Anamnese	Monolingual
94	DMar		u-me ((unverständlich, 1s)) e a	dar-me	assim tonturas pela cabeça. *I		Anamnese	Monolingual
94	DMar		eu médico lá de família acei eh	receitou-me	* * aconselhou-me para fazer te	interrogative	Anamnese	Monolingual
94	DMar		família acei/ eh receitou-me **	aconselho	para fazer terapia e fui fazer	interrogative	Anamnese	Monolingual
94	DMar		ara fazer terapia e fui fazer e	encontrei-	muito mal		Anamnese	Monolingual
94	DMar		á bastante tempo, mas o segundo	deu-me	cá de uma maneira. ((holt hörba		Anamnese	Monolingual
94	DMar		/ ao girar assim com o pescoco,	estalar-me	aqui muito. Foi por isso que o	declarative	Anamnese	Monolingual
94	DMar		nturas não. Deitada não, mas ao	por-me	a pé, vou já cair.		Anamnese	Monolingual
101	ILu		((blättert, 2 s)) Ora	deixe-me	cá ver. ((blättert, 3,5 s)) Ist	interrod...	Befund	Monolingual
101	ILu		Eu percebi,	estou-me	a rir, mas percebi.	interrogative	Befund	Monolingual
101	Umb		que me tinha acabado, não é, é	veio-me	e depois nunca mais me veio. Ac	exclamative	Befund	Monolingual
18	Mar		Ah, eu	sinto-me	bem	declarative	Aufnahme	Gedölmets
18	Mar		Sim *	Sinto-me	a rathar doente ((stottert 10s	other	Aufnahme	Gedölmets

Types: 106  
Tokens: 152  
Selected: 152  
Time: 32.14 s

Concordances

eigentlich  
Dolmetschen im Krankenhaus  
79 tokens  
1 types

det  
KS\_Toerring  
846 tokens  
1 types

myck, mycket  
KS\_Oeresund  
208 tokens  
2 types

ver-me, dizia-me, fui-me, Anda-me  
Dolmetschen im Krankenhaus  
152 tokens

Interactional context

((blättert, 2 s)) Ora **deixe-me** cá ver. ((blättert, 3,5 s)) Isto hoje foi complicado. ((blättert, 2 s))

Type: interrogative  
Diskursart[C]: Befund  
Konstellationstyp[C]: Monolingual

A [v] É? A senhora sabe algum? ((blättert, 2 s)) Ora **deixe-me** cá ver. ((blättert, 3,5 s)) Isto hoje foi complicado.

A [de] Ja? Wissen Sie welche? Also, lassen Sie mir sehen. Das heute war kompliziert.

P [v] É. Não é? É. Tá a ver?

P [de] Ja. Nicht wahr? Ja. Sehen Sie?

Partitur HTML

# Corpus query and analysis: EXAKT

The screenshot shows the EXAKT software interface. The main window displays a list of search results with columns for S, C..., Speaker, Left-Context, Match, Right-Context, and Type. A table of meta-data is highlighted in orange, showing 'Diskursart' and 'Konstellationstyp' for various results. A large white text overlay reads 'Meta Data (Situational context, ethnographic meta-data)'. The interface also shows a sidebar with corpus information and a bottom section with a transcription and a Partitur/HTML toggle.

S	C...	Speaker	Left-Context	Match	Right-Context	Type	Diskursar...	Konstellat...
91		SGa		E eu meti-me	no carro, feito maluco, vim soz		Aufklärung	Monolingual
92		INa	mentos. ((2s)) ((schmalz)) Eu,	falta-me	um exame. ((1,5s)) falta-me um	exclamati	Aufklärung	Monolingual
92		INa	Eu, falta-me um exame. ((1,5s))	falta-me	um exame para-lhe/* para *res	other	Aufklärung	Monolingual
92		DZé	*fiquei muito. ((1s)) Ainda me	deixe-me	andar. _E depois aquilo eh pas		Aufklärung	Monolingual
83		IFre	E eu quem sou?	Conhece-	a mim?	declarativ	Anamnese	Monolingual
94		DMar	((2s)) Olhe,	deu-me	((unverständlich, 1s)) e a dar-		Anamnese	Monolingual
94		DMar	u-me ((unverständlich, 1s)) e a	dar-me	assim fonturas pela cabeça. *I		Anamnese	Monolingual
94		DMar	eu médico lá de família ace/ eh	receitou-me	**aconselhou-me para fazer te	interrogat	Anamnese	Monolingual
94		DMar	família ace/ eh receitou-me **	aconselho	para fazer terapia e fui fazer	interrogat	Anamnese	Monolingual
94		DMar	ara fazer terapia e fui fazer e	encontrei-	muito mal		Anamnese	Monolingual
94		DMar	á bastante tempo, mas o segundo	deu-me	cá de uma maneira. ((hoff hörba		Anamnese	Monolingual
94		DMar	/ ao girar assim com o peçoço,	estalar-me	aqui muito. Foi por isso que o	declarativ	Anamnese	Monolingual
94		DMar	nturas não. Deitada não, mas ao	por-me	a pé, vou já cair.		Anamnese	Monolingual
101		ILu	((blättern, 2 s)) Ora	dexe-me	cá ver. ((blättern, 3,5 s)) Ist	interrod...	Befund	Monolingual
101		ILu		Eu percebi,	estou-me	a rir, mas percebi.	Befund	Monolingual
101		Umb	que me tinha acabado, não é, é	veio-me	e depois nunca mais me veio. Ac	exclamativ	Befund	Monolingual
18		Mar	Ah, eu	sinto-me	bem	declarativ	Befund	Monolingual
18		Mar	Sim *	Sinto-me	a calhar dnente ((stuttered 10s	other	Aufnahme ...	Gedolmets...

Meta Data  
(Situational context, ethnographic meta-data)



# Corpora

## Scandinavian semicommunication

- Recordings and transcriptions of inter-Nordic (Danish, Swedish, Norwegian) communication
  - radio broadcasts
  - school/university courses
  - conferences
- transcribed according to HIAT
- around 90 hrs of audio, 270,000 words

# Example

[28]

**[nsp]** ((0,6s))

**Dm1 [v]** ((holt Luft,0,5s)) Er det noget der kommer til udtryk sådan i dagligdagen, eller

**Sw9 [v]** så.

[29]

**Dm1 [v]** bare noget man kan se på medarbejderne? ((0,3s)) At danskerne er kraftigere, eller...

**Sw9 [v]**

Nej!

[30]

**[nsp]** ((0,5s))

((0,8s))

**Dm1 [v]** Nej.

**Sw9 [v]** Nej, nej. Det är de inte, inte så, men jag tror nog att vi påpekar för

[31]

**Sw9 [v]** varandra, "ät lite mer sallad" och ja...



# Corpora

## Interpreting in hospitals

- Recordings and transcriptions of multilingual doctor-patient communication mediated by non-trained interpreters
  - German/Turkish
  - German/Portuguese
  - German, Turkish and Portuguese control data
- transcribed according to HIAT
- around 25 hrs of audio, 170,000 words



# [www.exmaralda.org](http://www.exmaralda.org)

- Free download of software tools
- Documentation
- Example data
- (coming soon) Access to corpora
  
- Further demos on demand