



UNIVERSITÄT
LEIPZIG

Zugänge zu mündlichen Korpora für DaFZ aus dem Projekt ZuMult: Ansatzpunkte für Konstruktionsdidaktik und DDL- Aktivitäten

Christian Fandrych, Matthias Schwendemann, Franziska Wallner
Herder-Institut, Universität Leipzig

**Korpora, datengeleitetes Lernen (DDL) & Konstruktionen –
Anwendungsszenarien für die DaFZ-Vermittlungspraxis**

27.09.2024

Themen

1. Zugänge zu mündlichen Korpora für DaFZ aus dem Projekt ZuMult
2. Suche nach dem passenden Sprechereignis via ZuMal
3. Arbeit mit Transkripten via ZuViel
4. Wenn noch Zeit ist: Erstellung von Mikrohörübungen mit dem Tool ZuRecht

Das Projekt ZuMult

= Zugänge zu **multimodalen** Korpora gesprochener Sprache –
Vernetzung und zielgruppenspezifische Ausdifferenzierung

Förderung: DFG im Bereich „Wissenschaftliche
Literaturversorgungs- und Informationssysteme (LIS) -
Informationsinfrastrukturen für Forschungsdaten“

Laufzeit: 2018 – 2022

Beteiligte:

- Archiv für Gesprochenes Deutsch am IDS (Thomas Schmidt, Elena Frick)
- Hamburger Zentrum für Sprachkorpora (Kai Wörner)
- Herder-Institut Universität Leipzig (Christian Fandrych, Matthias Schwendemann, Franziska Wallner)



Webseite:

<https://zumult.org/>



Das Projekt ZuMult

Ziel: Verbesserung der Zugänglichkeit von digitalen mündlichen Sprachdaten für

1. Lehrende und Lernende im Kontext der Sprachvermittlung und
2. Forscher/-innen im Bereich Gesprächsanalyse und Variationslinguistik

vgl. Fandrych et al. (2023)



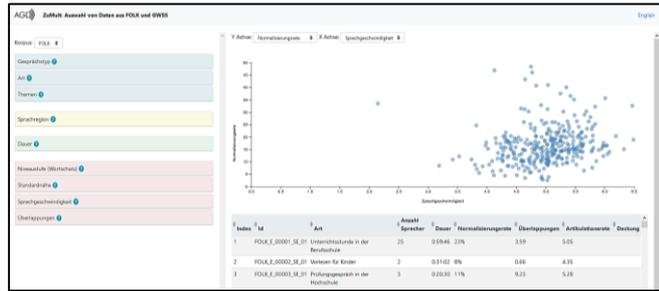
Webseite:

<https://zumult.org/>



WERKZEUGE AUS DEM PROJEKT ZUMULT FÜR DIE SPRACHDIDAKTIK

ZuMal Zugang zu Merkmalsauswahl von Sprechereignissen



**Nutzbar nach kostenfreier
Registrierung bei der DGD
(<https://dgd.ids-mannheim.de>)**

ZuViel Zugang zu Visualisierungselementen für Transkripte



<http://zumult.ids-mannheim.de/ProtoZumult/>

INFOS ZU DEN VIA ZUMULT NUTZBAREN KORPORA

Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK)



- umfasst große Bandbreite an privaten, institutionellen und öffentlichen Gesprächen aus dem gesamten deutschsprachigen Raum

Vergleichskorpus Gesprochene Wissenschaftssprache (GeWiss)

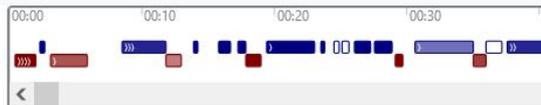


- umfasst Prüfungsgespräche, Expertenvorträge und studentische Vorträge auf Deutsch, Englisch, Polnisch und Italienisch

Ref Wortliste None

+	
+++	11
A	
ab	1
aber	14
abgeschlossen	1
abkühlen	1
ablaufen	1
abschneiden	1
abschöpfen	1
abstellen	2
abtrocknen	1
Acht	1
Achtung	2
ah	5
alle	6

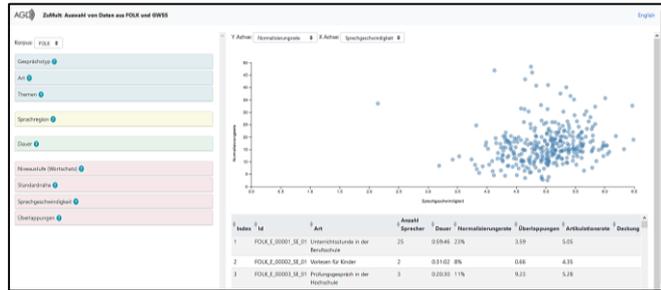
- 0001 (0.43)
- 0002 **CT** schnibbelst du schon mal die tomaten bitte
- 0003 (0.31)
- 0004 **AC** ja
- 0005 (0.44)
- 0006 **CT** und ich ähm schäl den knoblauch
- 0007 (2.62)
- 0008 **AC** schmeiß ich die tomaten dann gleich rein soll ich se halb schälen oder vierteln
- 0009 **CT** °h ich würd se vierteln
- 0010 (0.92)
- 0011 **AC** okay
- 0012 (1.58)
- 0013 **AC** ((stöhnt, 0.89s))
- 0014 (0.56)
- 0015 **AC** so
- 0016 **CT** und sonst wie war dein tag



Screenshot aus
dem ZuMult-Tool
ZuViel

Werkzeuge aus dem Projekt ZuMult für die Sprachdidaktik

ZuMal Zugang zu Merkmalsauswahl von Sprechereignissen



ZuViel Zugang zu Visualisierungselementen für Transkripte



Nutzbar nach kostenfreier
Registrierung bei der DGD
(<https://dgd.ids-mannheim.de>)

<http://zumult.ids-mannheim.de/ProtoZumult/>

Suche nach dem passenden Sprechereignis via ZuMal

AGD **ZuMult: Auswahl von Daten aus FOLK und GWSS**

Korpus: FOLK

Auswahl des Korpus

Gesprächstyp

Art

Themen

Sprachregion

Dauer

Niveaustufe (Wortschatz)

Standardnähe

Sprechgeschwindigkeit

Überlappungen

Wortarten

Mündlichkeitsphänomene

1

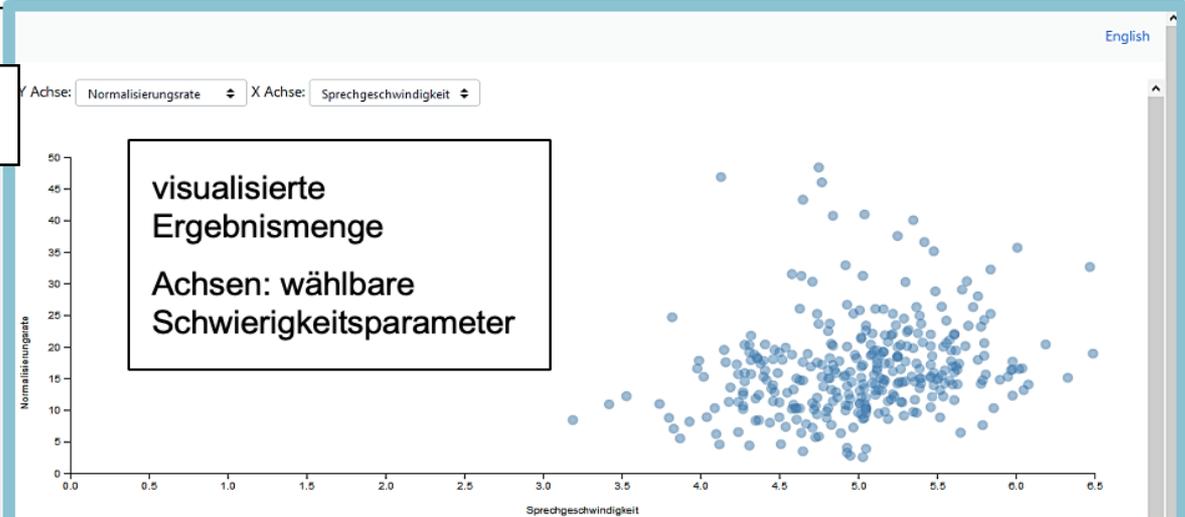
2

3

Y Achse: Normalisierungsrate X Achse: Sprechgeschwindigkeit

visualisierte Ergebnismenge

Achsen: wählbare Schwierigkeitsparameter



Index	Id	Art	Anzahl Sprecher	Dauer	Normalisierungsrate	Überlappungen	Artikulationsrate	Deckung
1	FOLK_E_00001_SE_01	Unterrichtsstunde in der Berufsschule	25	0:59:46	23%	3.59	5.05	
2	FOLK_E_00002_SE_01	Vorlesen für Kinder	2	0:31:02	8%	0.66	4.35	
3	FOLK_E_00003_SE_01	Prüfungsgespräch in der Hochschule	3	0:20:30	11%	9.23	5.28	
4	FOLK_E_00004_SE_01	Unterrichtsstunde in der Berufsschule	22					
5	FOLK_E_00005_SE_01	Unterrichtsstunde in der Berufsschule	22					

Ergebnisliste mit Werten für Schwierigkeitsparameter

Auswahlfiler

1) Metadaten

2) Schwierigkeitsparameter

3) Phänomendichte

Korpus: FOLK

Reset all

Y Achse: Normalisierungsrate

Gesprächstyp

Art

In dieses Feld können Suchbegriffe eingegeben werden.

Koch

Kochinteraktion (1)

Kommunikation beim Kochen (4)

Spielinteraktion zwischen Erwachsenen (9)

Spielinteraktion zwischen Erwachsenen (Videotelefonie) (1)

Spielinteraktion zwischen Erwachsenen unter Einsatz einer Spielekonsole (1)

Wortschatz (Niveaustufe)

Normalisierungsrate



Spalten auswählen

Id	Art
FOLK_E_00001_SE_01	Unterrichtsstu
FOLK_E_00002_SE_01	Vorlesen für K

Hier lassen sich weitere Spalten für die Ergebnisübersicht auswählen.

Spalten auswählen  ▾

◆ Index	◆ Id	◆ Art	◆ Dauer	◆ Normalisierungsrate	◆ Zu Viel Link
1	FOLK_E_00332_SE_01	Kochinteraktion	0:18:56	19%	→ZuViel
2	FOLK_E_00300_SE_01	Kommunikation beim Kochen	0:31:54	17%	→ZuViel
3	FOLK_E_00225_SE_01	Kommunikation beim Kochen	0:48:16	17%	→ZuViel
4	FOLK_E_00327_SE_01	Kommunikation beim Kochen	1:00:24	18%	→ZuViel
5	FOLK_E_00220_SE_01	Kommunikation beim Kochen	1:26:44	14%	→ZuViel

Die Sprechereignisse können über diese Links aufgerufen werden.

INFOS ZU DEN AUSWAHLFILTERN IN ZUMAL

Schwierigkeitsbezogene Parameter „Niveaustufe“

Korpus: FOLK ▾

Gesprächstyp ⓘ

Art ⓘ

Themen ⓘ

Sprachregion ⓘ

Dauer ⓘ

Niveaustufe (Wortschatz) ⓘ

Standardnähe ⓘ

Sprechgeschwindigkeit ⓘ

Überlappungen ⓘ

Niveaustufe (Wortschatz) ⓘ

Wortschatzliste: GOETHE_A1 ▾

Deckung in %: 80 ▴ ▾ - 100 ▴ ▾

0 20 40 60 80 100

ermöglicht eine wortschatzbezogene
Einstufung der einzelnen
Sprechereignisse nach dem
Sprachniveau

INFOS ZU DEN AUSWAHLFILTERN IN ZUMAL: Schwierigkeitsbezogene Parameter „Standardnähe“

Korpus: FOLK

Gesprächstyp

Art

Themen

Sprachregion

Dauer

Niveaustufe (Wortschatz)

Standardnähe

Sprechgeschwindigkeit

Überlappungen

Standardnähe ?

Anteil normalisierter Token in %: 0 - 10

0 20 40 60 80 100

Ermöglicht gezielte Suche nach Sprechereignissen, die besonders viele / wenige Abweichungen von der standardsprachlichen Lautung aufweisen

INFOS ZU DEN AUSWAHLFILTERN IN ZUMAL

Verfügbare Filter

➤ Nach Phänomendichte

Wortschatz (Niveaustufe)	?
Standardnähe (Normalisierungsrate)	?
Sprechgeschwindigkeit (Artikulationsrate)	?
Überlappungen	?
Wortarten	?
Mündlichkeitsphänomene	?

Mündlichkeitsphänomene

Häsitationen (NGHES)

Anteil Token in %: 0 - 11

äh, ähm, öh, hm ...

Niedriger Anteil Hoher Anteil

Ermöglicht gezielte Suche nach Sprechereignissen, die einen besonders hohen / niedrigen Anteil an ausgewählten Wortarten oder Mündlichkeitsphänomenen aufweisen

Für weitere ausführliche Infos zu den Filtern vgl. Fandrych et al. 2021, 2023

Ref Wortliste None

+

+++ 11

A 3

ab 1

aber 14

abgeschlossn 1

abkühlen 1

ablaufen 1

abschneiden 1

abschöpfen 1

abstellen 2

abtrocknen 1

Acht 1

Achtung 2

ah 5

alle 6



0001 (0.43)

0002 CT schnibbelst du schon mal die tomaten bitte

0003 (0.31)

0004 AC ja 1

0005 (0.44)

0006 CT und ich ähm schäl den knoblauch

0007 (2.62)

0008 AC schmeiß ich die tomaten dann gleich rein soll
ich se halb schälen oder vierteln

0009 CT °h ich würd se vierteln

0010 (0.92)

0011 AC okay

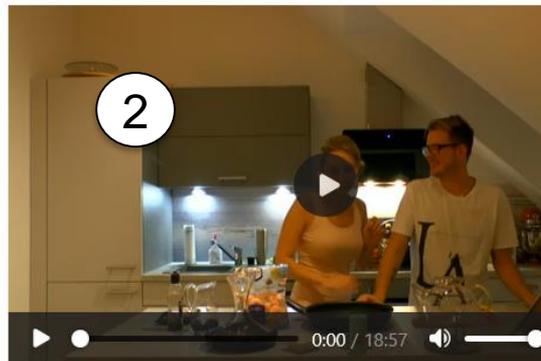
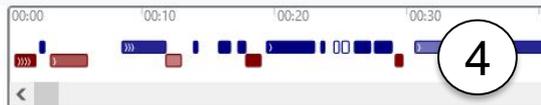
0012 (1.58)

0013 AC ((stöhnt, 0.89)

0014 (0.56)

0015 AC so

0016 CT und sonst wie war dein tag



1: Transkript
2: Video
3: Lemmaliste
4: Transkriptnavigator

Parameter	Sprechereignis (1)	Auswahl	pos=ART
0001	(0.43)		
0002	CT schnibbelst du schon mal die tomaten bitte		
0003	(0.31)		
0004	AC ja		
0005	(0.44)		
0006	CT und ich ähm schäl den knoblauch		
0007	(2.62)		
0008	AC schmeiß ich die tomaten dann gleich rein soll ich se halb schälen oder vierteln		
0009	CT °h ich würd se vierteln		
0010	(0.92)		
0011	AC okay		
0012	(1.58)		
0013	AC ((stöhnt, 0.89s)		
0014	(0.56)		
0015	AC so		
0016	CT und sonst wie war dein tag		

Parameter	Sprechereignis (1)	Auswahl	pos=ART
0001	(0.43)		
0002	CT schnibbelst Du schon mal die Tomaten bitte		
0003	(0.31)		
0004	AC ja		
0005	(0.44)		
0006	CT und ich <u>äh schäle</u> den Knoblauch		
0007	(2.62)		
0008	AC <u>schmeiße</u> ich die Tomaten dann gleich rein soll ich <u>sie</u> halb schälen oder vierteln		
0009	CT °h ich <u>würde sie</u> vierteln		
0010	(0.92)		
0011	AC okay		
0012	(1.58)		
0013	AC ((stöhnt, 0.89s)		
0014	(0.56)		
0015	AC so		
0016	CT und sonst wie war Dein Tag		

Vergleichen Sie die beiden Fassungen des Transkripts! Schreiben Sie in den Chat, welche Unterschiede es gibt?

Parameter	Sprechereignis (1)	Auswahl	pos=ART
0001	(0.43)		
0002	CT schnibbelst du schon mal die tomaten bitte		
0003	(0.31)		
0004	AC ja		
0005	(0.44)		
0006	CT und ich ähm schäl den knoblauch		
0007	(2.62)		
0008	AC schmeiß ich die tomaten dann gleich rein soll ich se halb schälen oder vierteln		
0009	CT °h ich würd se vierteln		
0010	(0.92)		
0011	AC okay		
0012	(1.58)		
0013	AC ((stöhnt, 0.89s))		
0014	(0.56)		
0015	AC so		
0016	CT und sonst wie war dein tag		

Aussprachenahes
Transkript

Parameter	Sprechereignis (1)	Auswahl	pos=ART
0001	(0.43)		
0002	CT schnibbelst Du schon mal die Tomaten bitte		
0003	(0.31)		
0004	AC ja		
0005	(0.44)		
0006	CT und ich <u>äh schäle</u> den Knoblauch		
0007	(2.62)		
0008	AC <u>schmeiße</u> ich die Tomaten dann gleich rein soll ich <u>sie</u> halb schälen oder vierteln		
0009	CT °h ich <u>würde sie</u> vierteln		
0010	(0.92)		
0011	AC okay		
0012	(1.58)		
0013	AC ((stöhnt, 0.89s))		
0014	(0.56)		
0015	AC so		
0016	CT und sonst wie war Dein Tag		

Standardorthografische Fassung
mit Markierung der vom
orthografischen Standard
abweichend realisierten
Einheiten

ZuViel: Zugang zu Visualisierungselementen für Transkripte

Parameter ✕

Linke Spalte (Wortliste) Rechte Spalte (Density und Audio/Video)

Transkript Normalisierte Formen ▼

Untertitel Transkribierte Formen ▼

Playback-Rate 100%

Transkript-Schriftart Sans Serif ▼ 16pt

Abweichungen zwischen transkribierter und normalisierter Form anzeigen

Sprechgeschwindigkeit im Transkript visualisieren

CT °h ich würde sie vierteln

CT schnibbelst Du schon mal die Tomaten bitte
(0.31)

AC ja
(0.44)

CT und ich äh schäle den Knoblauch

Optionen zur Transkriptanzeige in ZuViel

ZuViel: Zugang zu Visualisierungselementen für Transkripte

FOLK Kochinteraktion Parameter Sprechereignis (1) Auswahl pos=PTKMA

Ref Wortliste GOETHE_A2

0297 (2.38)

0298 AC die nudeln lassen wa **einfach**

0299 (0.28)

0300 CT hm

0301 AC hm

0302 (0.47)

0303 AC schmeck ich **einfach** ab (.) guck jetz nich auf die zeit

0304 (0.46)

0305 CT nee (.) des hab ich jetz auch nich (.) vor

0306 (0.56)

0307 AC ha

0308 CT hab ich au nich vorgehabt auf die zeit zu

0309 (1.41)

0310 AC könnten **ja** eigentlich au[ch]

0311 CT [a]chtung

0312 (1.04)

0313 AC jetz schon mal die spülmaschine räum

0314 (0.33)

00:00 00:10 00:20 00:30 00:40



Worum handelt es sich bei den markierten sprachlichen Einheiten? Antworten Sie im Chat!

19

ZuViel: Zugang zu Visualisierungselementen für Transkripte

The screenshot displays the FOLK Kochinteraktion interface. At the top, the title 'FOLK Kochinteraktion' and 'Parameter Sprechereignis (1) Auswahl' are visible. A search bar contains 'pos=PTKMA'. On the left, a 'Ref Wortliste' for 'GOETHE_A2' is shown with a list of words and their frequencies. The main transcript area shows a list of speech events with time stamps and text. Several words are highlighted in yellow, and modal particles are highlighted in red. A video player on the right shows a scene with two people in a kitchen.

Ref Wortliste: GOETHE_A2

- +++ 11
- A
- ✓ ab 1
- ✓ aber 14
- abgeschlossen 1
- abkühlen 1
- ablaufen 1
- abschneiden 1
- abschöpfen 1
- abstellen 2
- abtrocknen 1
- Acht 1
- ✓ Achtung 2
- ah 5
- ✓ alle 6

pos=PTKMA

0297 (2.38)

0298 AC die nudeln lassen wa einfach

0299 (0.28)

0300 CT hm

0301 AC hm

0302 (0.47)

0303 AC schmeck ich einfach ab (.) guck jetzt nicht auf die zeit

0304 (0.46)

0305 CT nee (.) des hab ich jetzt auch nicht (.) vor

0306 (0.56)

0307 AC ha

0308 CT hab ich au nicht vorgehabt auf die zeit zu schau

0309 (1.41)

0310 AC könnten ja eigentlich au[ch]

0311 CT [a]chtung

0312 (1.04)

0313 AC jetz schon mal die spülmaschine räum

0314 (0.33)

Transkriptausschnitt mit Umrahmung aller Modalpartikeln (PTKMA); gelb unterlegt sind alle Wörter bis A2

ZuViel: Zugang zu Visualisierungselementen für Transkripte

The screenshot shows the FOLK Kochinteraktion interface. At the top, there is a search bar with the text "pos=PTKMA" and a red box around it. Below the search bar, a modal window titled "Suchen und Hervorheben im Transkript" is open, displaying a list of grammatical categories. The "PTKMA" entry is highlighted with a red box. A red arrow points from the search bar to the "PTKMA" entry. On the right side of the interface, there is a video player showing a scene with two people in a kitchen. Below the video player, there is a red text box with the following text:

Transkriptausschnitt mit Umrahmung aller Modalpartikeln (PTKMA); gelb unterlegt sind alle Wörter bis A2

Category	Description	Examples
APPR	Präposition; Zirkumposition links	<i>in [der Stadt], ohne [mich]</i>
APPRART	Präposition mit Artikel	<i>im [Haus], zum [Beispiel]</i>
APZR	Zirkumposition rechts	<i>[von jetzt] an</i>
Interjektionen		
ITJ	Interjektion	<i>mhm, ach, tja</i>
Partikeln		
PTKA	Partikel bei Adjektiv oder Adverb	<i>am [schönsten], zu [schnell]</i>
PTKANT	Antwortpartikel	<i>ja, nein, danke, bitte</i>
PTKIFG	Intensitäts-, Fokus- und Gradpartikeln	<i>sehr, nur, ziemlich</i>
PTKMA	Modal- und Abtönungspartikeln	<i>halt, ja, schon</i>
PTKMWL	Partikel eines Mehrwortlexems	<i>[x] noch, immer [+gesteigertes Adjektiv], schon [+gesteigertes Adjektiv]</i>

0312 (1.04)

0313 **AC** jetz **schon** **mal** die spülmaschine räum

0314 (0.33)

ZuViel: Zugang zu Visualisierungselementen für Transkripte

The screenshot shows the FOLK Kochinteraktion interface. At the top, the search bar contains 'pos=PTKMA'. A red box highlights this search bar, with a red arrow pointing to a modal window titled 'Suchen und Hervorheben im Transkript'. The modal window lists search criteria: 'norm=haben', 'norm=nein', 'lemma=gehen', 'pos=VV.+', and 'pos=.+ .+'. Below the modal, a list of search results is shown, with 'ja' in result 0310 and 'jetz' in result 0313 highlighted with red boxes. On the left, a 'Ref Wortliste' dropdown is set to 'GOETHE_A2', and a table lists words like 'Acht', 'Achtung', 'ah', and 'alle'. On the right, a video player shows a scene with two people in a kitchen.

Suchen und Hervorheben im Transkript

- norm=haben (findet transkribierte Formen wie *ham*, *habn*)
- norm=nein (findet u.a. *nee*, *nö*, *na* ...)
- lemma=gehen (findet alle Flexionsformen wie *gehe*, *ging*, *gegangen*)
- pos=VV.+ (findet alle Vollverben)
- pos=.+ .+ (findet klitisierte Formen wie *weißte*, *brauchste*, *son*)

0307 AC ha

0308 CT hab ich au nich vorgehabt auf die zeit zu schau

0309 (1.41)

0310 AC könnten ja eigentlich au[ch]

0311 CT [a]chtung

0312 (1.04)

0313 AC jetz schon mal die spülmaschine räum

0314 (0.33)

Transkriptausschnitt mit Umrahmung aller Modalpartikeln (PTKMA); gelb unterlegt sind alle Wörter bis A2

AUSWAHL UND DOWNLOAD AUSGEWÄHLTER SEQUENZEN



Backen mit Freunden

FOLK_E_00372_SE_01_T_01

Parameters

Speech Event (1)

Selection



Ref wordlist

GOETHE_B1

+	
+++	19
A	
✓ ab	2
✓ aber	37
abgenutzt	1
abhauen	1
abmessen	1
abputzen	1
Abstellen	1
abtrocknen	3
✓ ach	5
✓ Achtung	5
achtzehnt	1
✓ Adresse	1
ah	28

0059 (0.6)

0060 SP so (,) als allererstes

Start Selection

End Selection

Annotations...

Partitur...

g hundertfufzig gramm weiche margarine hascht du (,) was zum abmesse

saage jo

0068 SP hundertfuchzig gramm weiche margarine

0069 (0.31)

0070 SP un drei eier

0071 (1.12)

0072 LH gibsch mir mol a dippe von do owe

0073 (1.76)

0074 RP °h der pinkne

0075 (0.36)

0076 SP ((lacht))

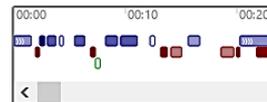
0077 LH nee das porzellanteil

0078 (1.22)

0079 SP °h awa wir müsse jetzt ah noch was zu esse bestelle ne

0080 (1.73)

pos=ART



AUSWAHL UND DOWNLOAD AUSGEWÄHLTER SEQUENZEN

FOLK Backen mit Freunden Parameters Speech Event (1) Selection

Ref wordlist: GOETHE_B1

+		19
+++		19
A		
✓ ab		2
✓ aber		37
abgenutzt		1
abhauen		1
abmessen		1
abputzen		1
Abstellen		1
abtrocknen		3
✓ ach		5
✓ Achtung		5
achtzehnt		1
✓ Adresse		1
ah		28

0059 (0.6)

0060 SP so (,) a

0061 (0.5)

0062 SP kommt

0063 LH ((schme

0064 SP ah voll

0065 (2.12)

0066 SP ähm

0067 (0.47)

0068 SP hundertfuchzig gramm weiche margarine

0069 (0.31)

0070 SP un drei eier

0071 (1.12)

0072 LH gibsch mir mol a dippe von do owe

0073 (1.76)

0074 RP *h der pinkne

0075 (0.36)

0076 SP ((lacht))

0077 LH nee das porzellanteil

0078 (1.22)

0079 SP *h awa wir müsse jetzt ah noch was zu esse bestelle ne

0080 (1.73)

hascht du (,) was zum abmesse

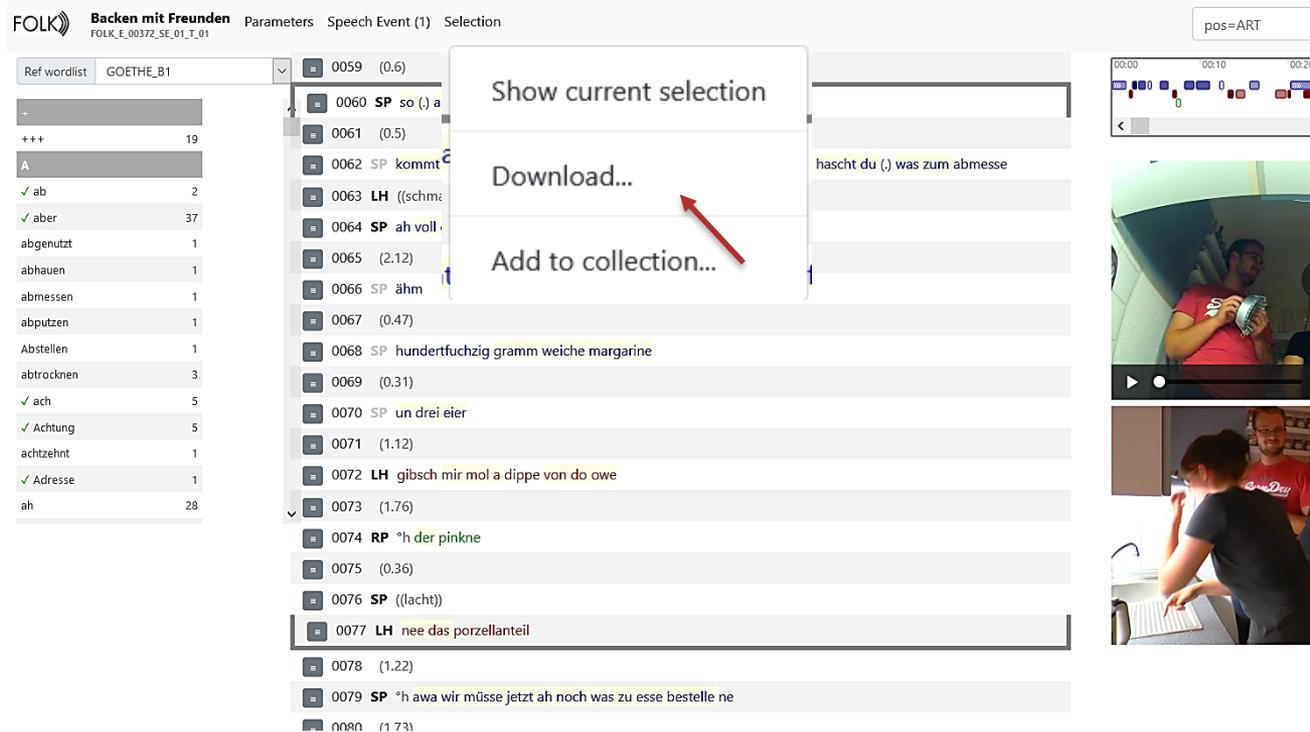
Show current selection

Download...

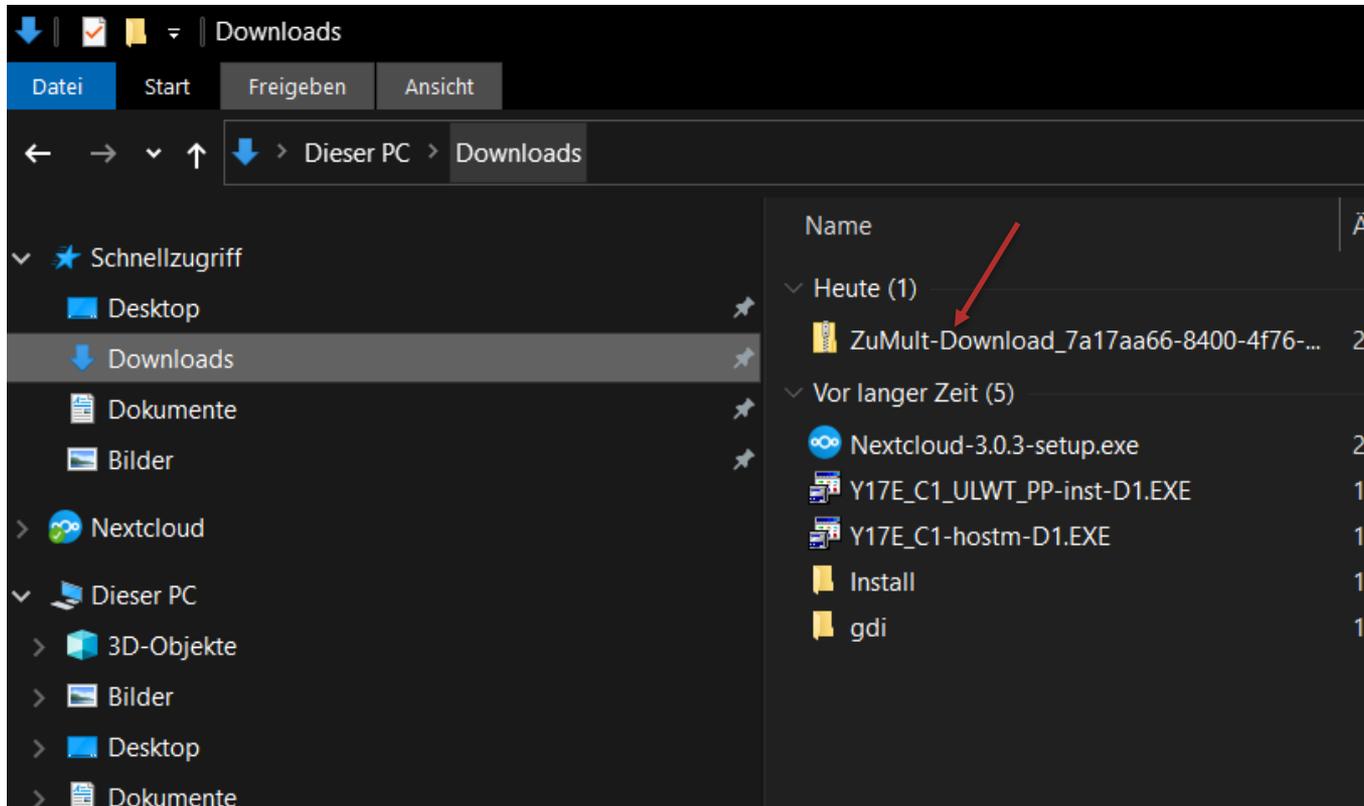
Add to collection...

pos=ART

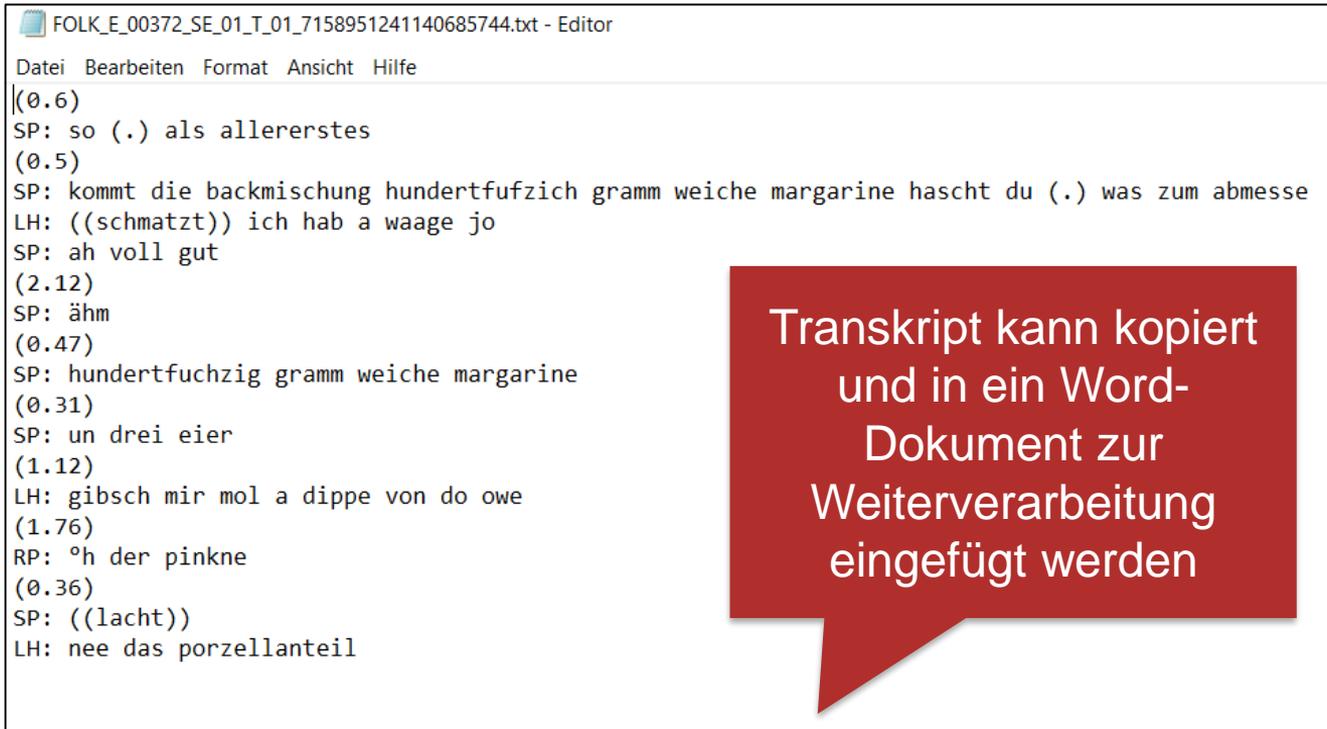
00:00 00:10 00:20



DOWNLOAD DURCH DOPPELKLICK ENTPACKEN



ANSICHT DES TRANSKRIPTS IM TEXTEDITOR



```
FOLK_E_00372_SE_01_T_01_7158951241140685744.txt - Editor
Datei Bearbeiten Format Ansicht Hilfe
(0.6)
SP: so (.) als allererstes
(0.5)
SP: kommt die backmischung hundertfufzich gramm weiche margarine hascht du (.) was zum abmesse
LH: ((schmatzt)) ich hab a waage jo
SP: ah voll gut
(2.12)
SP: ähm
(0.47)
SP: hundertfuchzig gramm weiche margarine
(0.31)
SP: un drei eier
(1.12)
LH: gibsch mir mol a dippe von do owe
(1.76)
RP: °h der pinkne
(0.36)
SP: ((lacht))
LH: nee das porzellanteil
```

Transkript kann kopiert
und in ein Word-
Dokument zur
Weiterverarbeitung
eingefügt werden

ZuViel: Zugang zu Visualisierungselementen für Transkripte

Optionen:

- **Markierung** im Transkript
 - Wortschatz innerhalb und außerhalb einer Niveaustufe
 - standardfern realisierte (normalisierte) Einheiten
- gezieltes **Ansteuern** relevanter Wortschatzeinheiten
 - Einheiten innerhalb oder außerhalb einer Niveaustufe (Wortliste)
- komprimierte **Navigation** im Sprechereignis
 - nach Sprecherwechseln
 - nach Sprechgeschwindigkeit
 - nach Standardnähe/-ferne
- Änderung der **Abspielgeschwindigkeit**
- **Download** von spezifischen Wortlisten zum Transkript

Didaktische Potenziale

- Binnendifferenzierung
- Vergleich von aussprachenaher Transkription und standardorthografischer Fassung
- Lenkung der Aufmerksamkeit auf ausgewählte sprachliche Einheiten
- Schneller Überblick
 - über unbekannte sprachliche Einheiten zur Vorentlastung
 - Gesprächsverlauf
- Grundlage zur Erstellung von Lückentexten und Mikrohörübungen

ein Nutzungsszenario



Öffnen Sie das Tool ZuMal und suchen Sie im FOLK nach einem geeigneten Sprechereignis! Notieren Sie im Chat, um was für eine Gesprächsart es sich handelt!

<https://zumult.ids-mannheim.de/ProtoZumult/prototype/dist/zuMal.jsp>

Korpus: FOLK Reset all

Gesprächstyp ?

Art ?

Themen ?

Sprachregion ?

Dauer ?

Wortschatz (Niveaustufe) ?

Standardnähe (Normalisierungsrate) ?

Sprechgeschwindigkeit (Artikulationsrate) ?

Überlappungen ?

Wortarten ?

Mündlichkeitsphänomene ? ←



Suchen Sie im FOLK nach einem geeigneten Sprechereignis! Notieren Sie im Chat, um was für eine Gesprächsart es sich handelt!

Spalten auswählen ⚙

Id	Art	Dauer	Normalisierungsrate	Zu Viel Link
FOLK_E_00001_SE_01	Unterrichtsstunde in der Berufsschule	0:59:45	22%	→ZuViel
FOLK_E_00002_SE_01	Vorlesen für Kinder	0:31:01	7%	→ZuViel
FOLK_E_00003_SE_01	Prüfungsgespräch in der Hochschule	0:20:29	9%	→ZuViel
FOLK_E_00004_SE_01	Unterrichtsstunde in der Berufsschule	1:26:47	23%	→ZuViel
FOLK_E_00005_SE_01	Unterrichtsstunde in der Berufsschule	1:15:31	28%	→ZuViel
FOLK_E_00006_SE_01	Unterrichtsstunde in der Berufsschule	0:41:39	26%	→ZuViel
FOLK_E_00007_SE_01	Unterrichtsstunde in der Berufsschule	1:28:23	21%	→ZuViel
FOLK_E_00008_SE_01	Unterrichtsstunde in der Berufsschule	0:56:14	23%	→ZuViel

Wortschatz (Niveaustufe) ?**Standardnähe (Normalisierungsrate)** ?**Sprechgeschwindigkeit (Artikulationsrate)** ?**Überlappungen** ?**Wortarten** ?**Mündlichkeitsphänomene** ?

Modalpartikeln (PTKMA)

Anteil Token in %: - 

Niedriger Anteil

Hoher Anteil

[Reset](#)

Ich suche ein Gespräch, in dem ganz viele Modalpartikeln vorkommen.



Suchen Sie im FOLK nach einem geeigneten Sprechereignis! Notieren Sie im Chat, um was für eine Gesprächsart es sich handelt!

Spalten auswählen ⚙

Index	Id	Art	Dauer	Normalisierungsrate	Anteil PTKMA	Zu Viel Link
1	FOLK_E_00439_SE_01	Telefongespräch	0:50:56	10%	5.01%	→Zu Viel



Ref Wortliste None ▾

+

+++

18

A

ab

1

aber

39

abgeben

1

absolut

2

abspielen

1

ach

2

achtundfünfzig

1

achtzehn

1

Achtzigerjahr

2

Adapter

2

Ahnung

2

alle

4

allein

1

0001 (0.24)

0002 **HB** hi ich glaub wir sind jetzt verbunden

0003 (0.5)

0004 **HB** °hh (.) h°0005 **LF** jo sind wir h° (.) hi hanne[s]0006 **HB** °h0007 **HB** [ja]°hh h° ja h°0008 **LF** °hh h°0009 **HB** °hh schön dass wir dann

0010 (0.25)

0011 **HB** mal nach langer zeit wieder zusammenfinden

0012 (0.3)

0013 **HB** is ja [auch]0014 **LF** [ja]

0015 (0.72)

0016 **HB** n paar

0017 (0.43)

0018 **HB** tage wieder her dass man das letzte mal gesprochen hat

Lassen Sie sich alle Abweichungen zwischen der transkribierten und der orthographisch normalisierten Form anzeigen! Welches Wort ist das erste, das im Transkript hervorgehoben wird?

Ref Wortliste: None

Parameter

Linke Spalte (Wortliste) Rechte Spalte (Density und Audio/Video)

Transkript Transkribierte Formen

Playback-Rate 100%

Transkript-Schrif Sans Serif 12pt

Abweichungen zwischen transkribierter und normalisierter Form anzeigen

Sprechgeschwindigkeit im Transkript visualisieren



Lassen Sie sich alle Abweichungen zwischen der transkribierten und der orthographisch normalisierten Form anzeigen! Welches Wort ist das erste, das im Transkript hervorgehoben wird?

allein

1

0015 (0.72)

 0016 **HB** n paar

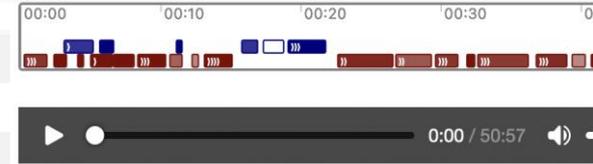
0017 (0.43)

 0018 **HB** tage wieder her dass man das letzte mal gesprochen hat



Ref Wortliste None ▾

		0001 (0.24)
+		0002 HB hi ich <u>glaub</u> wir sind jetzt verbunden
+++	18	0003 (0.5)
A		0004 HB °hh (.) h°
ab	1	0005 LF <u>jo</u> sind wir h° (.) hi hanne[s]
aber	39	0006 HB °h
abgeben	1	0007 HB [ja]°hh h° ja h°
absolut	2	0008 LF °hh h°
abspielen	1	0009 HB °hh schön dass wir dann
ach	2	0010 (0.25)
achtundfünfzig	1	0011 HB mal nach langer zeit wieder zusammenfinden ne
achtzehn	1	0012 (0.3)
Achtzigerjahr	2	0013 HB is ja [auch]
Adapter	2	0014 LF [ja]
Ahnung	2	0015 (0.72)
alle	4	0016 HB n paar
allein	1	0017 (0.43)
		0018 HB tage wieder her dass man das letzte mal gesprochen hat



Lassen Sie sich alle Abweichungen zwischen der transkribierten und der orthographisch normalisierten Form anzeigen! Welches Wort ist das erste, das im Transkript hervorgehoben wird?



Ref Wortliste None

+	
+++	18
A	
ab	1
aber	39
abgeben	1
absolut	2
abspielen	1
ach	2
achtundfünfzig	1
achtzehn	1
Achtzigerjahr	2
Adapter	2
Ahnung	2
alle	4
allein	1

- ☰ 0001 (0.24)
- ☰ 0002 **HB** hi ich glaub wir sind jetzt verbunden
- ☰ 0003 (0.5)
- ☰ 0004 **HB** °hh (.) h°
- ☰ 0005 **LF** jo sind wir h° (.) hi hanne[s]
- ☰ 0006 **HB** °h
- ☰ 0007 **HB** [ja]°hh h° ja h°
- ☰ 0008 **LF** °hh h°
- ☰ 0009 **HB** °hh schön dass wir dann
- ☰ 0010 (0.25)
- ☰ 0011 **HB** mal nach langer zeit wieder zusammenfinden ne
- ☰ 0012 (0.3)
- ☰ 0013 **HB** is ja [auch]
- ☰ 0014 **LF** [ja]
- ☰ 0015 (0.72)
- ☰ 0016 **HB** n paar
- ☰ 0017 (0.43)
- ☰ 0018 **HB** tage wieder her dass man das letzte mal gesprochen hat



Markieren Sie nun alle Modalpartikeln im Transkript. Welche sind die ersten 2 und auf welcher Zeile sind Sie zu finden?



Ref Wortliste: None ▾

Suchen und Hervorheben im Transkript



APPR

Präposition; Zirkumposition links

in [der Stadt], ohne [mich]

APPRART

Präposition mit Artikel

im [Haus], zum [Beispiel]

APZR

Zirkumposition rechts

[von jetzt] an

Interjektionen

ITJ

Interjektion

mhm, ach, tja

Partikeln

PTKA

Partikel bei Adjektiv oder Adverb

am [schönsten], zu [schnell]

PTKANT

Antwortpartikel

ja, nein, danke, bitte

PTKIFG

Intensitäts-, Fokus- und Gradpartikeln

sehr, nur, ziemlich

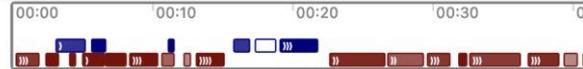
PTKMA

Modal- und Abtönungspartikeln

halt, ja, schon

PTKMWL

Partikel eines Mehrwortlexems

[x] noch, immer [+gesteigertes Adjektiv], schon [+gesteigertes Adjektiv]

Markieren Sie alle Modalpartikeln im Transkript. Welche sind die ersten 2 und auf welcher Zeile sind Sie zu finden?

allein

1

0015 (0.72)

0016 HB n paar

0017 (0.43)

0018 HB tage wieder her dass man das letzte mal gesprochen hat



Ref Wortliste: None ▾

	0013	HB	is ja [auch]
	0014	LF	[ja]
+			
+++	18		
A			
ab	1	0016	HB n paar
aber	39	0017	(0.43)
abgeben	1	0018	HB tage wieder her dass man das letzte mal gesprochen hat
absolut	2	0019	(0.69)
abspielen	1	0020	LF ja so sieht s aus
ach	2	0021	(0.41)
achtundfünfzig	1	0022	LF °h (.) joa h°
achtzehn	1	0023	(0.32)
Achtzigerjahr	2	0024	LF wie war so wie war so deine letzte woche was hast du so gemacht
Adapter	2	0025	(0.87)
Ahnung	2	0026	HB äh ich war bei meinem bruder (.) und der hatte (.) ne seminararbeit
alle	4	0027	(0.2)
allein	1	0028	HB fertigzustellen da hab ich ihm (.) bei geholfen
		0029	(0.29)
		0030	HB war halt n bisschen stressig weil der



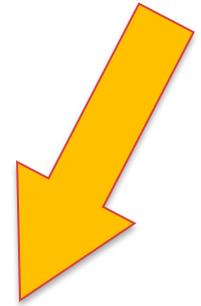
Markieren Sie alle Modalpartikeln im Transkript. Welche sind die ersten 2 und auf welcher Zeile sind Sie zu finden?

Infos zur Nutzung von ZuMal & ZuViel

auf dem FaDaF-Youtube Channel:



https://www.youtube.com/watch?v=X_wj4Uchb3U



Fandrych, C. / Schwendemann, M. & Wallner, F. (2021): „Ich brauch da dringend ein passendes Beispiel ...“: Sprachdidaktisch orientierte Zugriffsmöglichkeiten auf Korpora der gesprochenen Sprache aus dem Projekt ZuMult. In: Informationen Deutsch als Fremdsprache 50: 6, 711–729.

Zugänge zu mündlichen Korpora für DaFZ aus dem Projekt ZuMult: Ansatzpunkte für Konstruktionsdidaktik und DDL-Aktivitäten

Ressourcen / Tools

GeWiss: Gesprochene Wissenschaftssprache kontrastiv; online zugänglich unter <https://gewiss.uni-leipzig.de/> sowie über die Datenbank für Gesprochenes Deutsch (DGD) unter <https://dgd.ids-mannheim.de> (26.09.2024)

FOLK: Forschungs- und Lehrkorpus gesprochenes Deutsch, zugänglich über die Datenbank für Gesprochenes Deutsch (DGD) unter <https://dgd.ids-mannheim.de> (26.09.2024)

ZuMal: Zugang zu Merkmalsauswahl von Sprechereignissen (Prototyp), online unter: <https://zumult.ids-mannheim.de/ProtoZumult/index.jsp> (26.09.2024)

ZuViel: Zugang zu Visualisierungs-Elementen für Transkripte (Prototyp), online unter: <https://zumult.ids-mannheim.de/ProtoZumult/index.jsp> (26.09.2024)

ZuRecht: Zugang zu Recherchen auf Transkripten (Prototyp), online unter: <https://zumult.ids-mannheim.de/ProtoZumult/index.jsp> (26.09.2024)

Literatur

Fandrych, Christian/Schwendemann, Matthias/Wallner, Franziska (2021): „Ich brauch da dringend ein passendes Beispiel ...“: Sprachdidaktisch orientierte Zugriffsmöglichkeiten auf Korpora der gesprochenen Sprache aus dem Projekt ZuMult. In: InfoDaF 50, 6, 711–729.

Fandrych, Christian/Schmidt, Thomas/Wallner, Franziska/Wörner, Kai (2023a): Zugänge zu multimodalen Korpora gesprochener Sprache für DaF und DaZ. Themenheft zu Korpora Deutsch als Fremdsprache (KorDaF) 3:1, Online unter: <https://kordaf.tujournals.ulb.tu-darmstadt.de/issue/92/info/> (26.09.2024)

Fandrych, Christian/Meißner, Cordula/Schwendemann, Matthias/Wallner, Franziska (2023b): ZuMal: Zielgruppenspezifische Gesprächsauswahl aus Korpora gesprochener Sprache. In: Korpora Deutsch als Fremdsprache 3(1), 13–43, DOI: <https://doi.org/10.48694/kordaf.3725> (26.09.2024)

Frick, Elena/Helmer, Henrike/Wallner, Franziska (2023): Zugang zur Recherche in Transkripten mit dem Tool ZuRecht. Nutzungsmöglichkeiten für Forschung und Sprachvermittlung. In: Korpora Deutsch als Fremdsprache (KorDaF) 3:1., 44–71, DOI: <https://doi.org/10.48694/kordaf.3730> (26.09.2024)

Schmidt, Thomas, (2023): FOLK - Das Forschungs- und Lehrkorpus für Gesprochenes Deutsch, Korpora Deutsch als Fremdsprache 3(1), 166–169, DOI: <https://doi.org/10.48694/kordaf.3737> (26.09.2024)

Schmidt, Thomas / Schwendemann, Matthias / Wallner, Franziska (2023): ZuViel: Transkriptvisualisierung und Arbeiten mit Transkripten. In: Korpora Deutsch als Fremdsprache 3(1), 72–91, DOI: <https://doi.org/10.48694/kordaf.3723> (26.09.2024)