

TM

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Institut für
Romanische
Philologie
& ITG

Vortrag Erlangen
18.6.2012

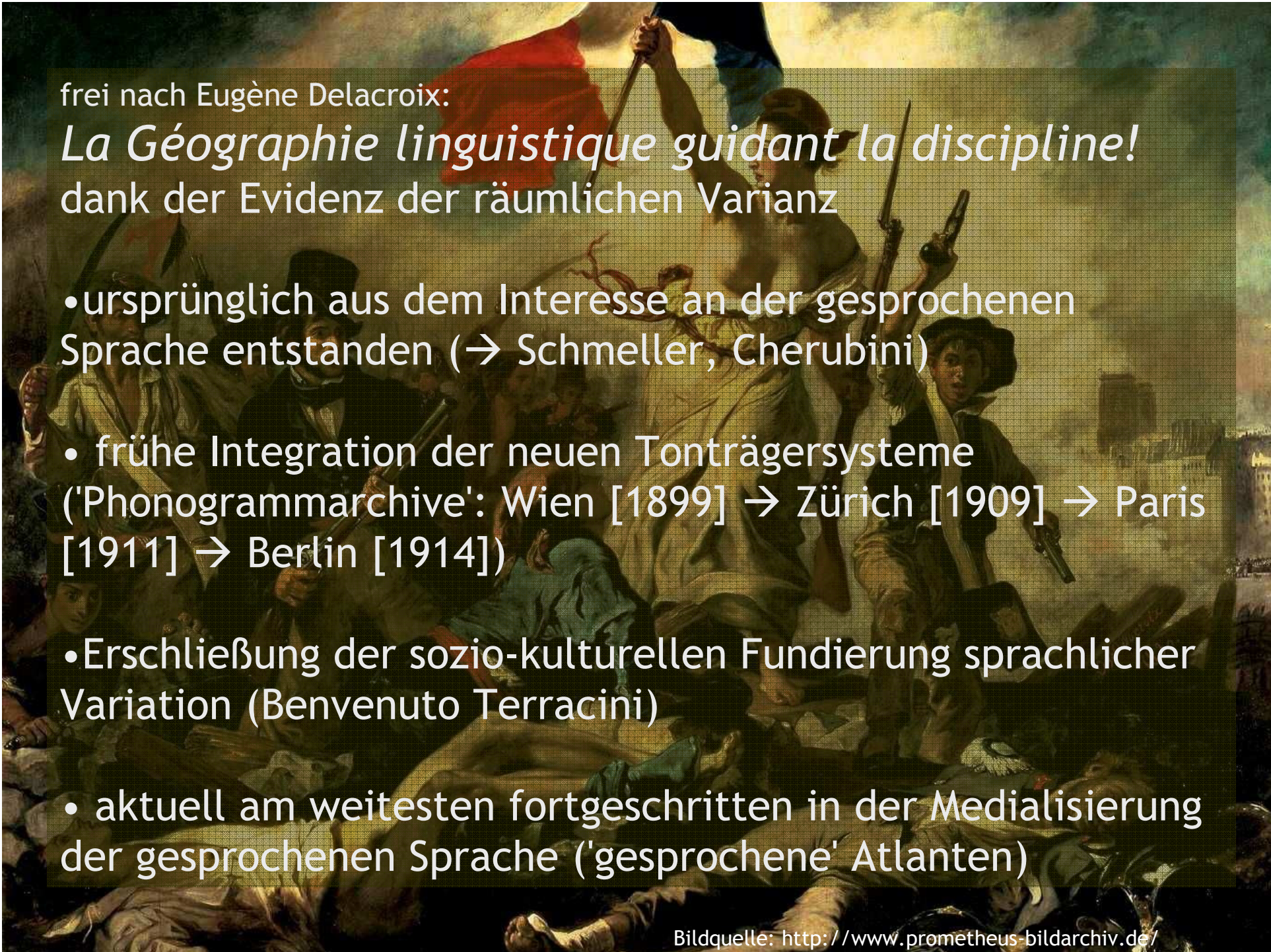
GeoLing2.0 - Aktuelle Projekte zur webbasierten Sprachgeographie

Ein Blick in die Werkstatt

Thomas Krefeld & Stephan Lücke

Inhalt

1. Ein Blick in die Wissenschaftsgeschichte vor dem Hintergrund der Neuen Medien
2. Unsere Projekte
3. Aspekte der Digitalisierung
4. Daten und Metadaten
5. Analyse und Visualisierung
6. Web 2.0 - Kollaboratives Arbeiten



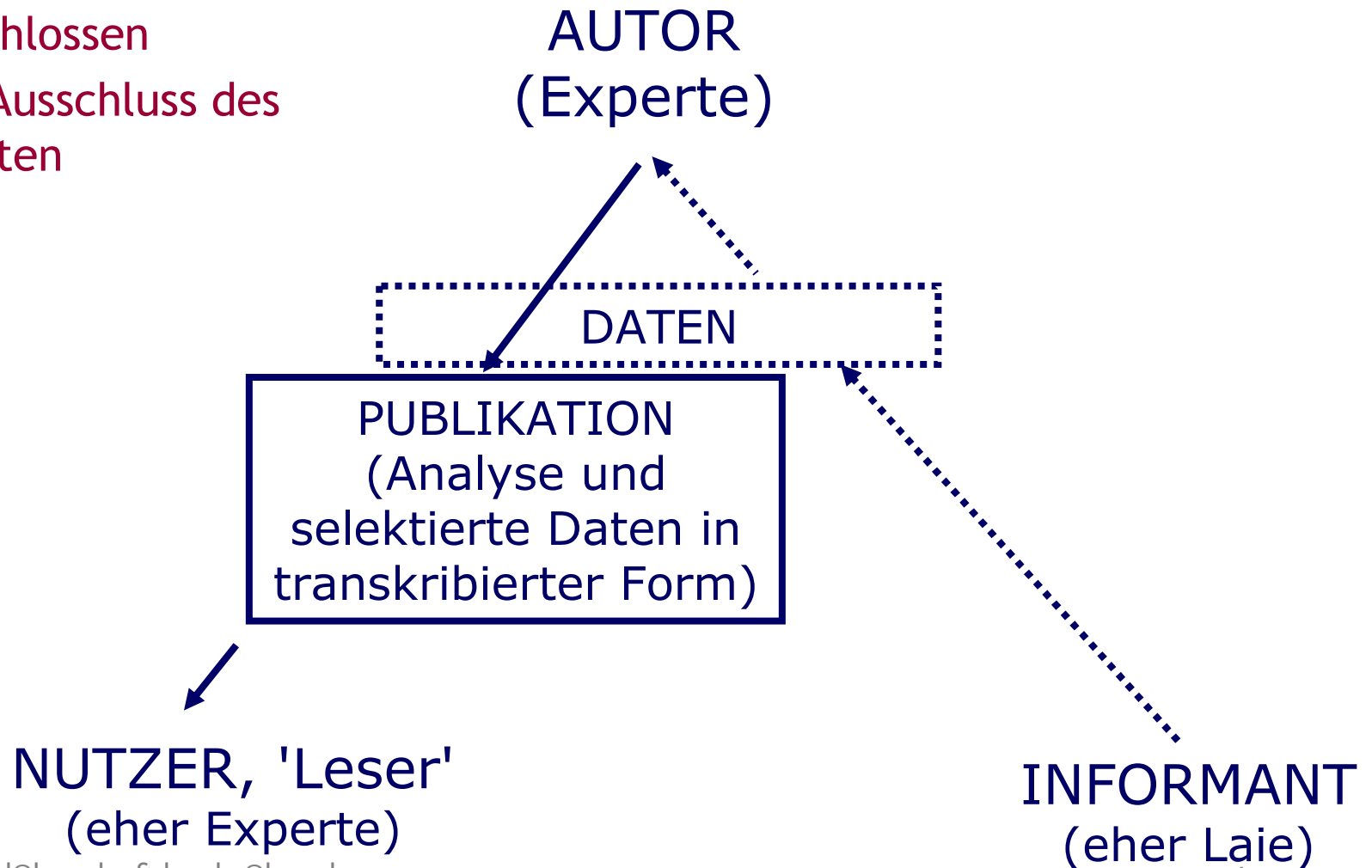
frei nach Eugène Delacroix:

La Géographie linguistique guidant la discipline!
dank der Evidenz der räumlichen Varianz

- ursprünglich aus dem Interesse an der gesprochenen Sprache entstanden (→ Schmeller, Cherubini)
- frühe Integration der neuen Tonträgersysteme ('Phonogrammarchiv': Wien [1899] → Zürich [1909] → Paris [1911] → Berlin [1914])
- Erschließung der sozio-kulturellen Fundierung sprachlicher Variation (Benvenuto Terracini)
- aktuell am weitesten fortgeschritten in der Medialisierung der gesprochenen Sprache ('gesprochene' Atlanten)

traditionelle sprachwissenschaftliche Kommunikation (1)

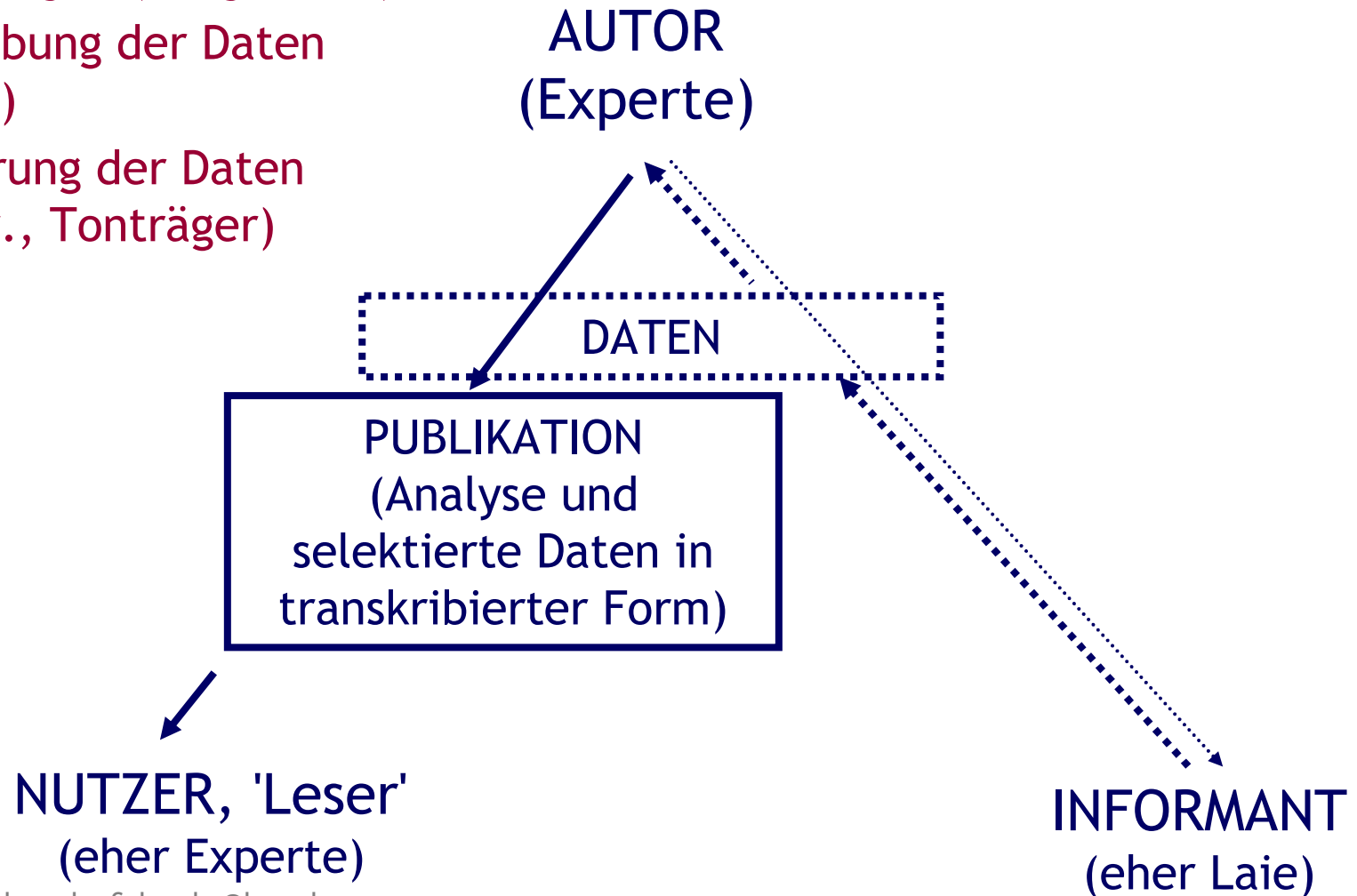
- unidirektional
- abgeschlossen
- unter Ausschluss des Informanten



traditionelle sprachwissenschaftliche Kommunikation (2)

Rolle der Medien: Werkzeuge des Autors mit dem Ziel der selbständigen ('dinglichen') Publikation

- ev. Erhebung der Daten (Schrift!)
- Speicherung der Daten (Transkr., Tonträger)
- Analyse
- Druck

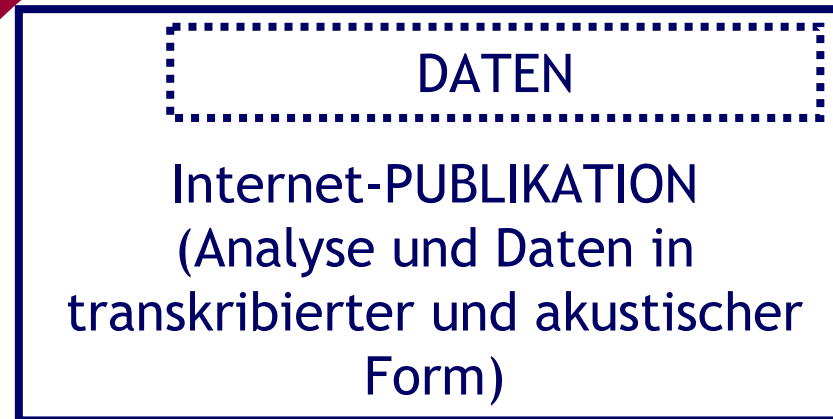


sprachwissenschaftliche Kommunikation nach der medialen Revolution

Integration von
Dokumentation und
Analyse
→ Öffnung des
Korpus zur
freien
Recherche

AUTOR
(Experte)

radikale Steigerung
der Authentizität der
Daten und des
Informanten
(bleibt als
Individuum
erkennbar)



NUTZER, 'Leser'
(eher Experte)

INFORMANT
(eher Laie)

allerdings noch in einer Übergangsphase

Phasen des Übergangs

- Digitalisierung von Papier
- mehr oder weniger starke Nutzung der neuen medialen Möglichkeiten: z.B. die so genannten 'sprechenden' Atlanten

historischer Pionier: ALD von Hans Goebel)

- <http://ald.sbg.ac.at/ald/ald-i/index.php?id=m002&lang=de>

Werner König (Augsburg, 2006-2008)

- <http://sprachatlas.bayerische-landesbibliothek-online.de/>

(Belege für die online-Version nacherhoben und mit den Karten des *Kleinen Bayerischen Sprachatlas* verknüpft)

- mehr oder weniger starke Integration interaktiver und kollaborativer Optionen

Inhalt

1. Ein Blick in die Wissenschaftsgeschichte vor dem Hintergrund der Neuen Medien
2. Unsere Projekte

3. Aspekte der Digitalisierung
4. Daten und Metadaten
5. Analyse und Visualisierung
6. Web 2.0 - Kollaboratives Arbeiten



Steckbrief AsiCa

- Kalabresischer Dialekt
- Audioaufnahmen, großteils phonetisch transkribiert
- Spontanmaterial und gelenktes Material
- Rund 400.000 Tokens
- Datenerhebung in 2004
- Onlineportal seit 2006
- DFG-Förderung 2004-2006
- www.asica.gwi.uni-muenchen.de



Steckbrief ASD

- „Audioatlas Siebenbürgisch-sächsischer Dialekte“
- Über 350 Stunden Audio-Material
- Teils gelenkt (Wenkersätze), überwiegend spontan
- Ursprünglich auf Tonband
- 2212 Dateien, 141 GB wav, 11 GB mp3
- Aufnahmen zw. 1968 und 1973
- Aktuell gefördert vom BKM (bis 2014)
- www.asd.gwi.uni-muenchen.de



Steckbrief AdIS

- “Atlante linguistico digitale dell'Italia e della Svizzera meridionale”
- Tiefendigitalisierung des „Sprach- und Sachatlas Italiens und der Südschweiz“ von Jaberg/Jud (1928-40)
- Gelenkt erhobenes Material
- Datenanalyse und Erzeugung von online-Karten
- Kombination mit außersprachlichen Faktoren
- Startphase, derzeit rund 40 Karten erfaßt
- www.adis.gwi.uni-muenchen.de



Steckbrief Metropolitalia

- Innovatives Projekt im web 2.0 (social software)
- Online-Portal in Form eines Spiels (GWAP) zur Sammlung von Sprachdaten
- Observatorium der räumlichen Variation des Italienischen (pluridimensional)
- Aktuell DFG-gefördert
- beta.metropolitalia.org

Inhalt

1. Ein Blick in die Wissenschaftsgeschichte vor dem Hintergrund der Neuen Medien
 2. Unsere Projekte
 3. Aspekte der Digitalisierung
-
4. Daten und Metadaten
 5. Analyse und Visualisierung
 6. Web 2.0 - Kollaboratives Arbeiten

Verschiedene Phasen der Digitalisierung:

- Digitalisierung als graduelles Phänomen
- „oberflächliche“ D. → 10011011
- „digitale Tiefenerschließung“ (OCR, manuell)
- Strukturierung (automatisch)

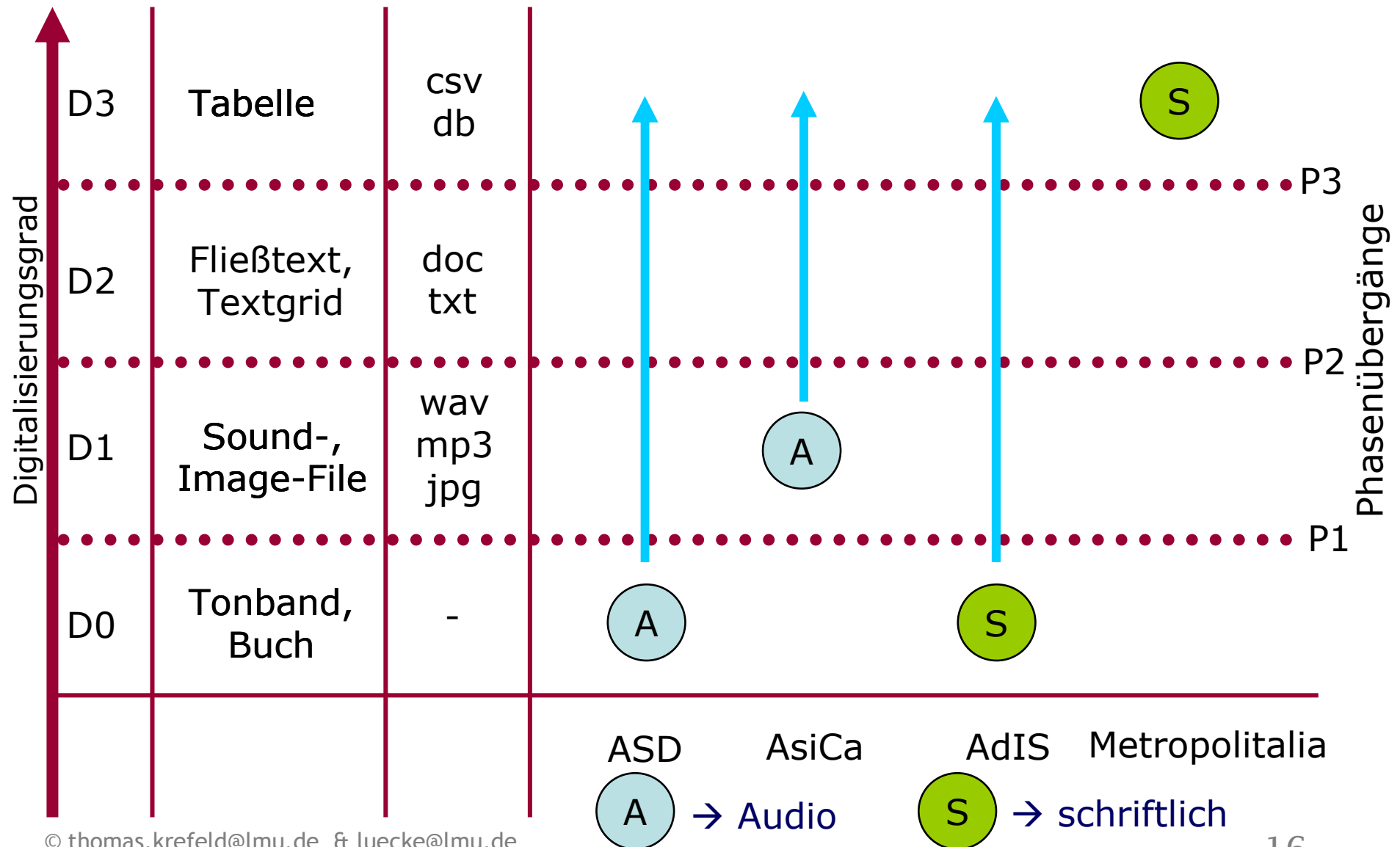
Beispiel:

Buchseite → Scanner → Bild → OCR → Text →
Strukturierung → Datenbank (Tabelle)

Unterschiedlicher Digitalisierungsgrad

- ASD: Tonbandaufnahmen aus den 60er/70er Jahren des 20.Jh.s
- ASiCa: Audiodaten auf Minidisk von 2004
- AdIS: gedruckte Sprachkarten
- Metropolitalia: dynamisches Datenerfassungssystem, vollständig digital

Ausgangslage bei Projektbeginn und Entwicklung



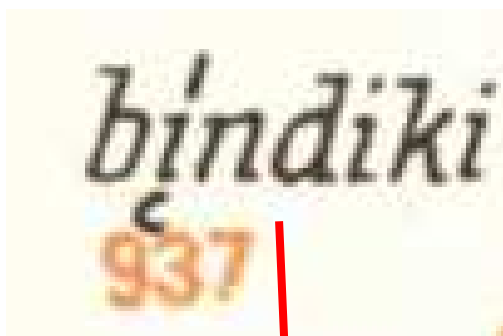
Digitalisierung schriftlichen Materials (Beispiel AdIS)

Status D0



Jaberg, Karl / Jud, Jakob (1928-1940): Sprach- und Sachatlas Italiens und der Suedschweiz, 8 Bde.,; Index 1960

Digitalisierung schriftlichen Materials (Beispiel AdIS): Zeichenkodierung



bindiki
937

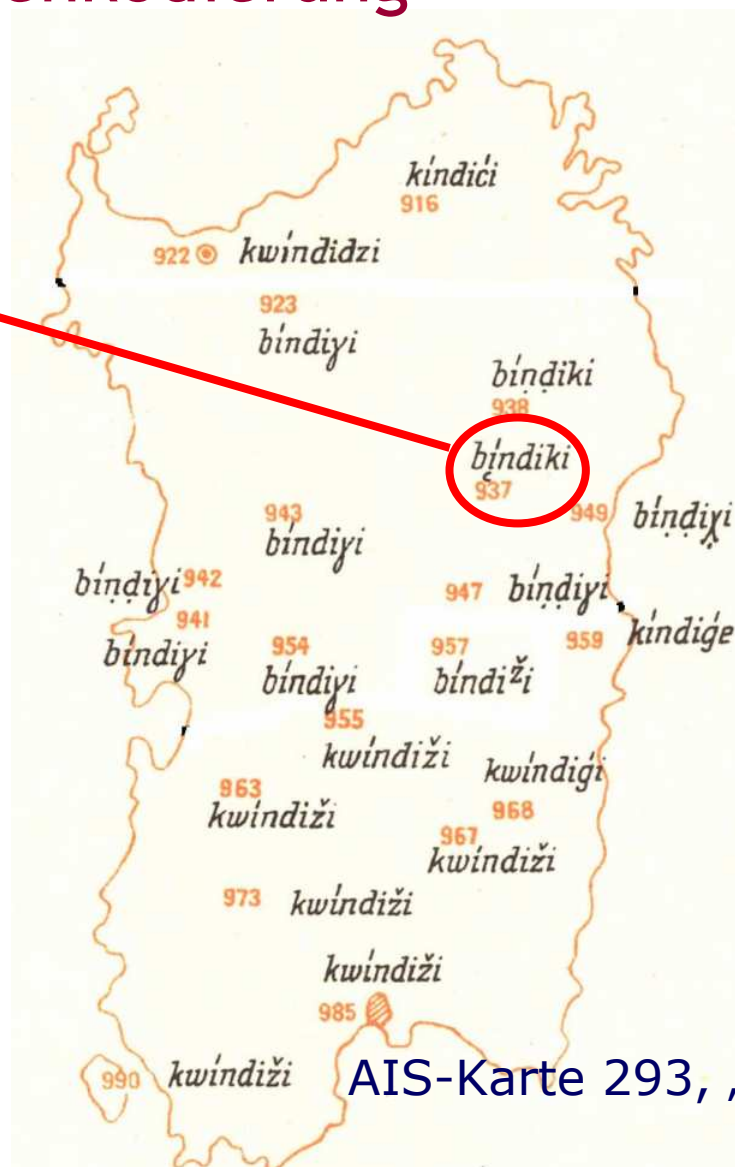
Regeln I: "proprietär"

bi(/ndiki

Datenbank

Regeln II: IPA

[*'bindiki*]



AIS-Karte 293, „quindici“

Digitalisierung schriftlichen Materials (Beispiel AdIS): Zeichenkodierung



Folgende Basiszeichen, die nicht in der ASCII-Tabelle enthalten sind, werden wie folgt transkribiert:

α	Griechisches Alpha	a1
β	Griechisches Beta	b1
γ	Griechisches Gamma	g1
δ	Griechisches Delta	d1
θ	Griechisches Theta	t1
φ	Griechisches Phi	f1
χ	Griechisches Chi	c1
ȳ	Zeichen für velares „n“ (dt. kling)	n1
œ	Ligatur œ	o1
ə	Schwa	e1 (wie z.B. „e“ in dt. „Tanne“)
ĭ	Durchgestrichenes „l“	l1 (wie z.B. in ital. „figlio“)

Diakritika werden grundsätzlich hinter das Basiszeichen gesetzt, dem sie zugeordnet sind.

Diakritika:

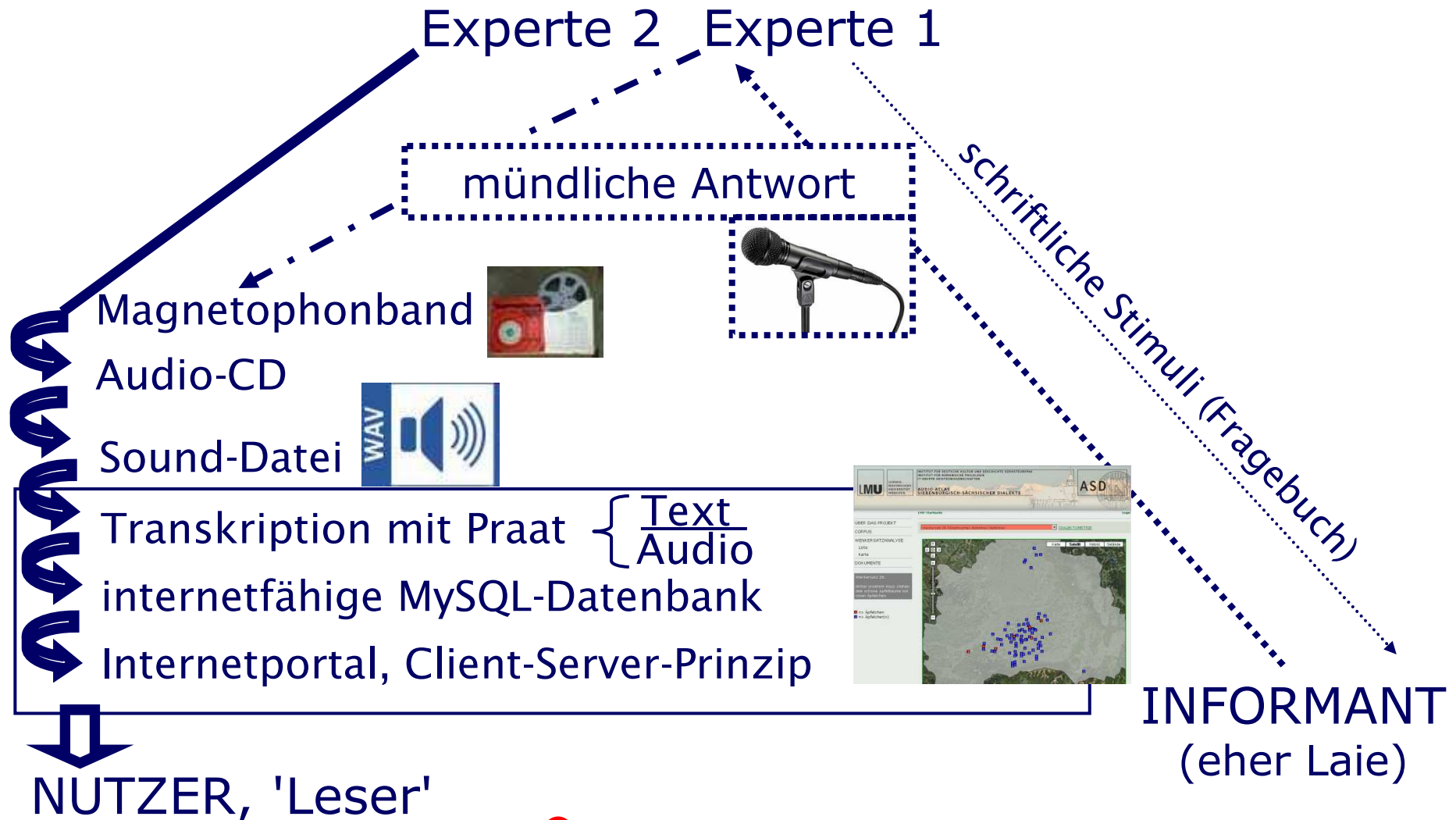
ä	:	Zwei Punkte über Basiszeichen	Doppelpunkt: (auch „ä“, „ö“, „ü“!)
ø	?	Punkt unter Basiszeichen	(nachgestelltes) Fragezeichen: ?
ø	(	Nach rechts geöffnetes Häkchen unter Basiszeichen	offene Klammer: (
ʳ	)	Nach links öffnender Halbkreis (spiritus lenis) über Basiszeichen	geschlossene Klammer:)
ċ	/	Akut auf Basiszeichen	Slash: /
ā	-	Waagerechter Strich über Basiszeichen	Minuszeichen: -
ã	~	Tilde auf Basiszeichen	Tilde: ~
buon ^a	0	Hochgestellter Buchstabe	[Buchstabe]0 (Beispiel: buon ^a -> buona0)
š	^	Hacek	^ (Beispiel: s^)
š	§	Nach unten öffnender Häkchen	§

Transkriptionsregeln

http://www.metropolitalia.gwi.uni-muenchen.de/dokumente_AIS/Transkriptionsregeln.doc

Digitalisierung von Audiomaterial (Beispiel ASD)

1



Digitalisierung Phasenübergang P3 (Fließtext → Tabelle)

```
intervals [184]:  
  xmin = 537.75340126847767  
  xmax = 540.8524228312366  
  text = "sugnu \t.troppu stanku u ne ne\sh\shu stasira"
```

id_wort	wort	id_form	Interview	quest	Wort_nr	Text_nr	Intervall	Subintervall	position	Sprecher
3443	sugnu	4445	Bag1mDQ1	F54:	784	205	184	1	1	l:
3444	troppu	1272	Bag1mDQ1	F54:	785	205	184	1	2	l:
3445	stanku	4367	Bag1mDQ1	F54:	786	205	184	1	3	l:
3446	u	5170	Bag1mDQ1	F54:	787	205	184	1	4	l:
3447	me	34496	Bag1mDQ1	F54:	788	205	184	1	5	l:
3448	neffu	22399	Bag1mDQ1	F54:	789	205	184	1	6	l:
3449	stasira	4376	Bag1mDQ1	F54:	790	205	184	1	7	l:

Referenzsystem

Inhalt

1. Ein Blick in die Wissenschaftsgeschichte vor dem Hintergrund der Neuen Medien
2. Unsere Projekte
3. Aspekte der Digitalisierung
4. Daten und Metadaten

5. Analyse und Visualisierung
6. Web 2.0 - Kollaboratives Arbeiten

CARTINA INTERATTIVA

SCANSIONI AIS

DOCUMENTI

BANCA DATI

- => tipo la mascella (129)
- => tipo la faccia/la vista (75)
- => tipo la ganassa (74)
- => tipo la guancia (46)
- => tipo la gota (35)
- => tipo gargia (16)
- => tipo la drempa (10)
- => tipo la barra (6)
- => tipo la kavanu (3)

In latino classico la guancia ('parte laterale della faccia tra lo zigomo e la mandibola') era chiamata BŪCCA; questa parola si è bene conservata in parecchie lingue romanze, ma i riflessi hanno tutti un altro significato dovuto a uno spostamento metonimico, quello dell'it. *bocca*, fr. *bouche* ecc.: 'cavità nella parte inferiore della faccia umana, che serve per la respirazione, la nutrizione e la fonazione'. Sembra dunque che BŪCCA abbia sostituito già in epoca latina la parola ŌS 'bocca', perdendo nello stesso tempo il suo significato originale. (Eventualmente la scomparsa di ŌS 'bocca' era favoreggiata dalla somiglianza fonetica con il lat. ŌSSUM 'osso'.) La carta presente, comunque, documenta tutta una serie di parole di origini diverse che designano la parte laterale della faccia. [Thomas Krefeld]

lessico => la guancia



tipo la ganassa - commento

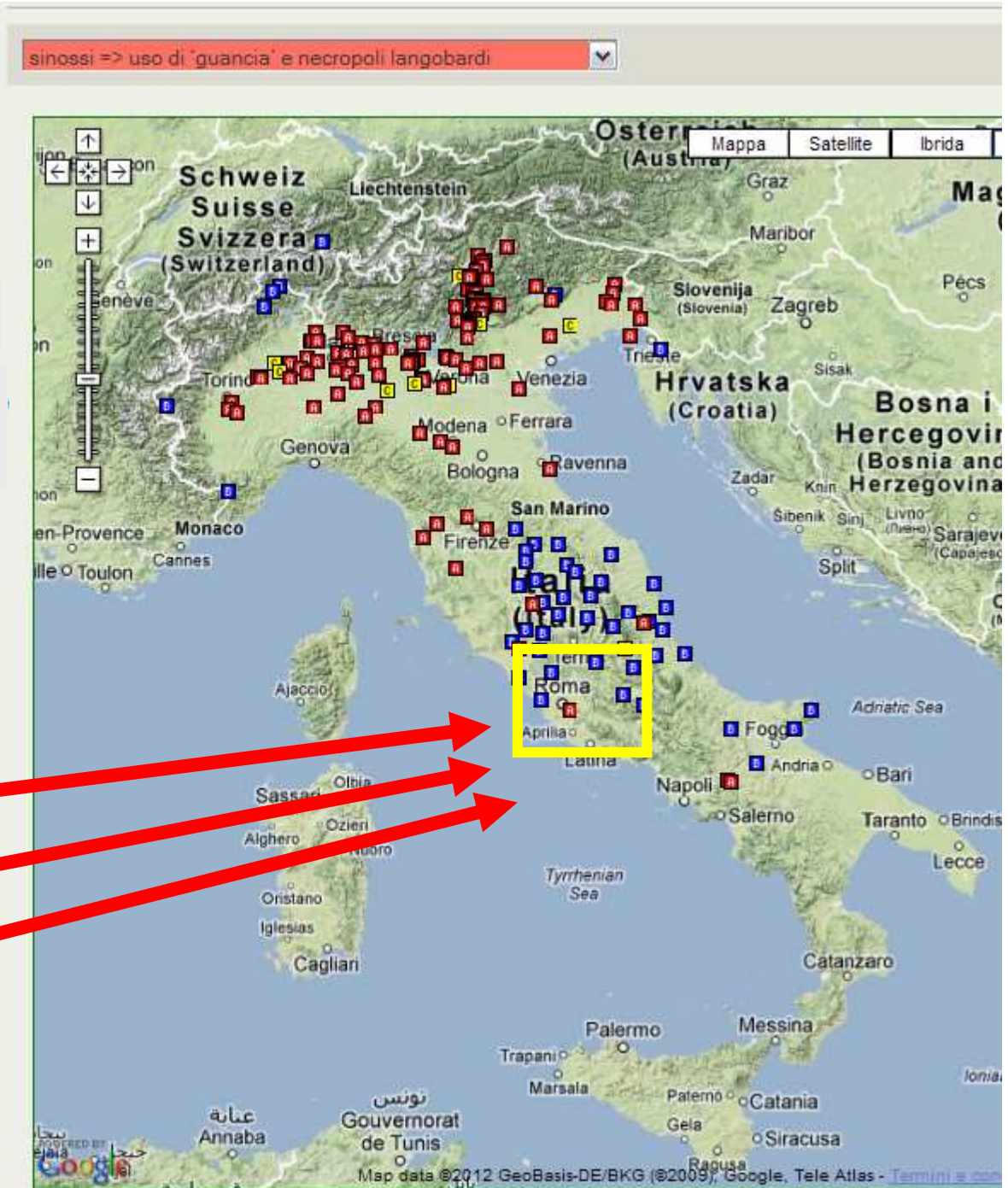
Deriva dal greco *gnatos*. Questo prestito era in opposizione della parola latina MAXILLA per evitare una polisemia tra i significati *la ganascia* e *la guancia*, cionondimeno anche *gnatos* abbia sofferto lo stesso cambio nel significato come MAXILLA. Il prestito è stato iniziato dall'Esarcato d'Italia, un territorio sotto l'amministrazione bizantina tra il VI - l'VIII secolo con il capoluogo Ravenna (Blasco Ferrer, 1994, 136). [Schwemmlein, Miriam]

Metadaten 1: Kommentare

Kombination sprachlicher mit außersprachlichen Realitäten

- R => langob. Nekropolen
(120)
- B => *guancia* 'Wange' in
ital. Dialekten
(49)
- i
- G => ital. Ortsnamen auf
-engo, -ingo
(10)

Metadaten 2:
Geschichte
Infrastruktur
Demographie



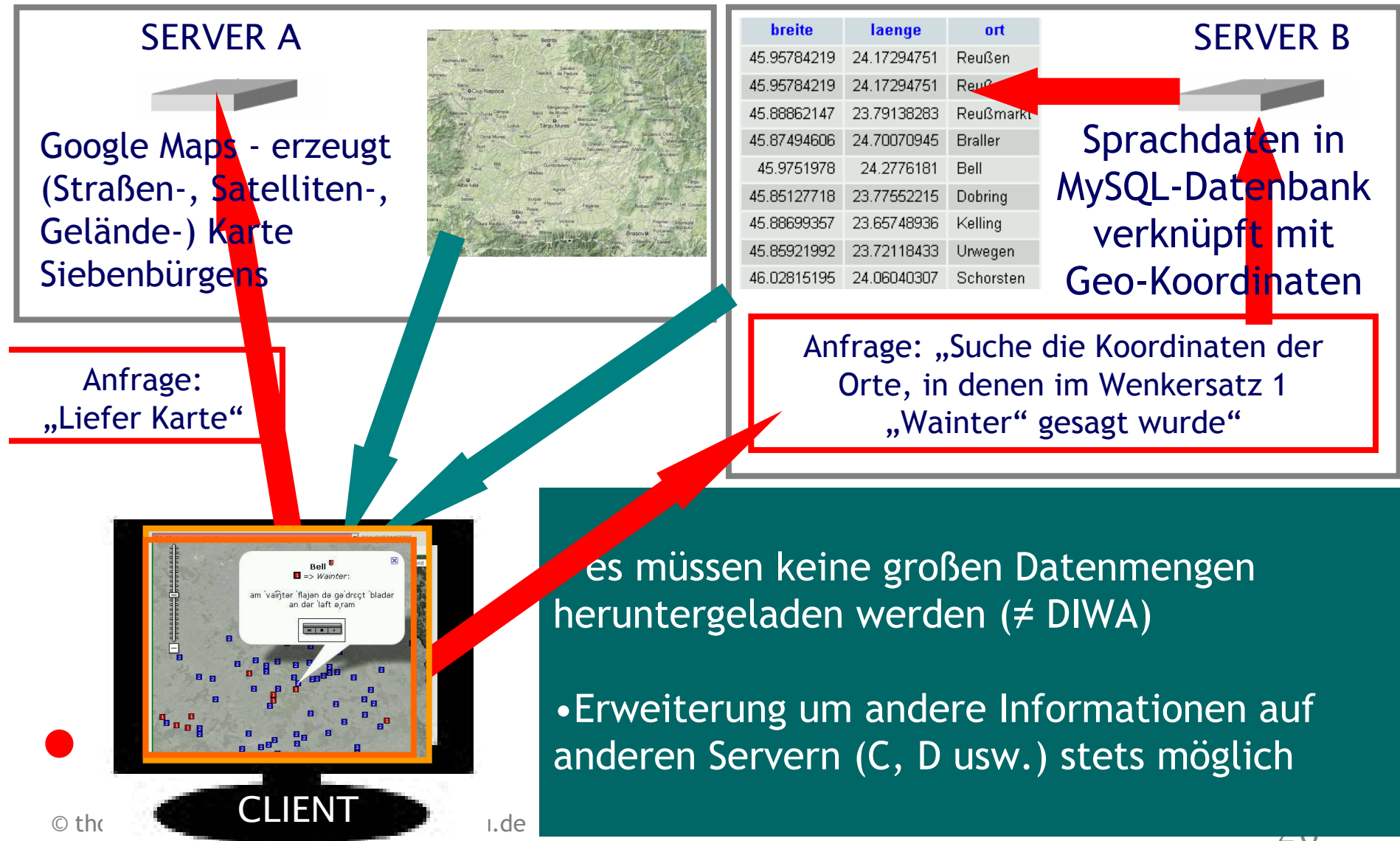


Metadaten 3: bibliographische Informationen



Metadaten n: usw.

Kombination von Daten und Metadaten: Georeferenzierung im Detail



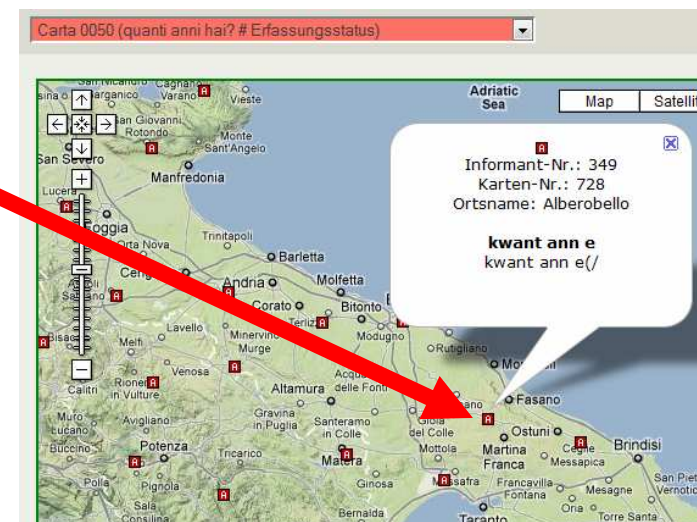
Beispiel: Suche nach Informanten, die beim Stimulus „quanti anni hai“ den Anlaut „kwa“ verwendet haben, und Generierung der Karte

SQL: select id_informanti_ais, ais_karte, aeusserung from aeusserungen_ais where ais_karte = 50 and aeusserung like 'kwa%'

id_informanti_ais	Fremdschlüssel	ais_karte	Nummer der ais_karte	aeusserung
350		50		kwa/nt a/nni0 tye-/ni0
349		50		kwant ann e(/
348		50		kwand a/nne1 ti-/ne1
id_informanti_ais	id_comune	Herkunft des Informanten		
349	316	5697		
349	317	5700		
349	348	8092		
349	350	6120		

Georeferenzierung

id_comune	name	longitude	latitude
5697	Iglistra	13.2556	42.0703
5700	Trasacco	13.5354	41.9579
6051	Alberobello	17.238	40.7869
6120	Carovigno	17.6558	40.709
8092	Spinazzola	16.0924	40.9681



wünschenswerte Optionen, die
durch interessierte Nutzer ergänzt
werden können
Google-Earth-Layer:

Optionen
umgesetzt
im Projekt ASD

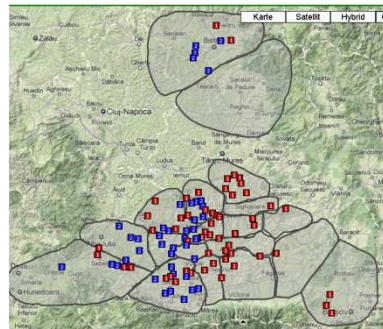
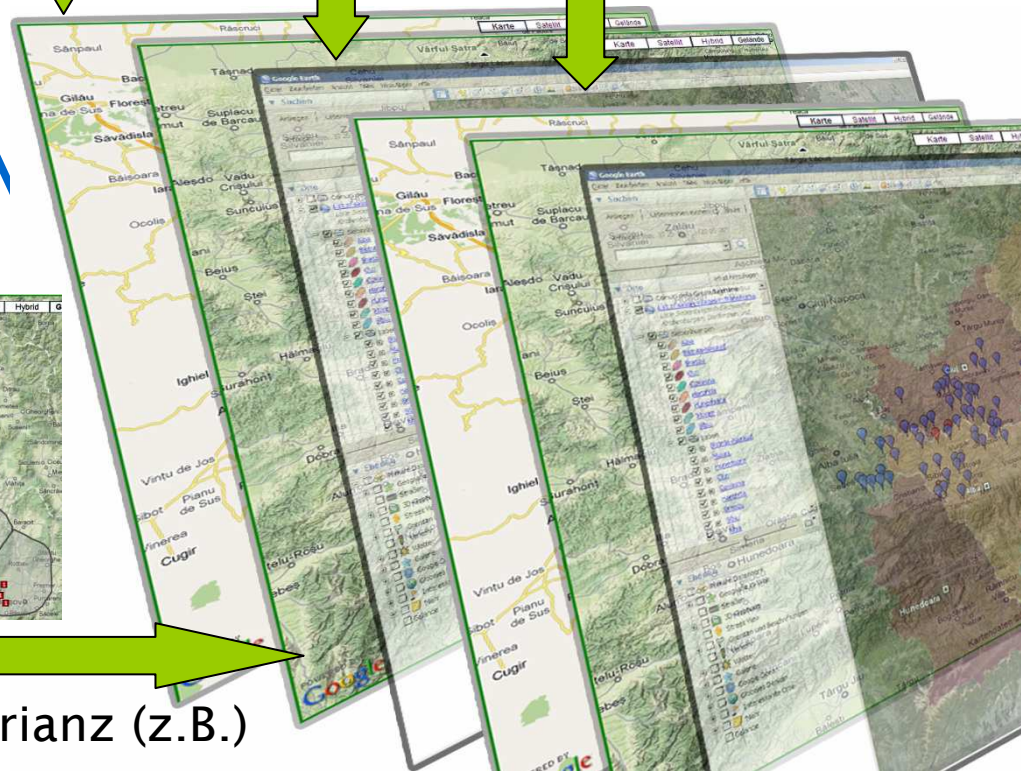


moderne Infrastruktur (z.B. Flughäfen)

demographische Entwicklung

historische Karten

etc.



Corpus-
Datenbank

Stärke der Varianz (z.B.)

Inhalt

1. Ein Blick in die Wissenschaftsgeschichte vor dem Hintergrund der Neuen Medien
 2. Unsere Projekte
 3. Aspekte der Digitalisierung
 4. Daten und Metadaten
 5. Analyse und Visualisierung
-
6. Web 2.0 - Kollaboratives Arbeiten

Visualisierung: Liste (und Konkordanzen)

ÜBER DAS PROJEKT

CORPUS

WENKERSATZANALYSE

Liste

Karte

Wenkersatz **26**: Hinter unserem Haus stehen drei schöne Apfelbäume mit roten Äpfelchen

Audio	Ort (Interview-Nr.)	Mundartbezirk	Wenkersatz
	Abtsdorf bei Agnetheln (1113-05)	4	hoĩndr a:zəm 'haĩs ʃtɛ:ən dro: 'hiʃ 'a:pəlbeĩm mæt 'riūd̩n ,a:pəltʃɛrn
	Abtsdorf bei Marktschelken (1077a)	2	hoĩnj̥dər eāzər 'ʃtɔf ʃtəʊn t̩sviĩn hiʃh 'a:pəlbiĩm mæt 'hiʃ 'riūt ,a:pəltʃɛr
	Agnetheln (894-04)	17	hoĩnj̥dər a:zəm 'haĩs ʃtiō:n t̩svəĩ hi:ʃ 'a:pəlbaōmtçər mæt 'ruĩdn ,a:pəltçɛrən
	Almen (1004)	4	hoĩnj̥dər a:zəm 'he:s ʃtən t̩svəĩ
	Almen (1004)	4	'hiʃ 'a:pəlbĩmtçər mæt 'riūd̩n 'a:pəltçən
	Almen (1005a)	4	hoĩnj̥dər a:zəm 'høūs ʃtəʊn t̩svəĩ 'hoĩj 'a:pəlboūmtçər mæt 'riūd̩n 'a:pəltçɛrn

[...]



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

INSTITUT FÜR DEUTSCHE KULTUR UND GESCHICHTE SÜDOSTEUROPAS
INSTITUT FÜR ROMANISCHE PHILOLOGIE
IT-GRUPPE GEISTESWISSENSCHAFTEN

AUDIO-ATLAS
SIEBENBÜRGISCH-SÄCHSISCHER DIALEKTE

ASD



LMU-Startseite

Login

ÜBER DAS PROJEKT

CORPUS

WENKERSATZANALYSE

DOKUMENTE

Datenbestand: 2212 | Anzahl gefundene Dateien: 18

[Alle Dateien](#) | [Nur transkribierte Dateien](#) | [Alle Filter entfernen](#) | [Karte bei Google Earth](#)

Audio	Ort	Jahr	Alter	Corpus	Stichwörter
Filter	Heltau	Filter	Filter	Filter	Filter
	Heltau (S)		60	Bukarest	Gedicht (?)
	Heltau (S)		32	Bukarest	Erinnerungen
	Heltau (S)		63	Bukarest	Spezialitäten
	Heltau (S)		14	Bukarest	Märchen vom Rotkäppchen
	Heltau (S)		13	Bukarest	Fasching
	Heltau (S)		14	Bukarest	Erlebnis
	Heltau (S)		63	Bukarest	Landwirtschaft; Markt; Maschinen
	Heltau (S)		60	Bukarest	kirchlicher Frauenverein
	Heltau (S)		65	Bukarest	Gemeinde
	Heltau (S)	1968	17	Bukarest	Schulreise
	Heltau (S)	1968	17	Bukarest	Ausflüge; Ferienarbeit als Tagelöhner; Ferienpläne
	Heltau (S)	1968	24	Bukarest	Theatergruppe; Inhalt eines Theaterstücks
	Heltau (S)	1968	60	Bukarest	Heimindustrie aus der Sicht der Frauen; Alteisensammler
	Heltau (S)	1968	60	Bukarest	Hochzeit
	Heltau (S)	1968		Bukarest	Wenkersätze
	Heltau (S)	1968	75	Bukarest	Wollweberei früher; Weberschule; Rolle der Frauen in der Weberei und in der Wirtschaft; Jagd; Heltauer Ackerbürger; Mithilfe der Kinder
	Heltau (S)		63	Bukarest	Gedichte
	Heltau (S)		60	Bukarest	Wenkersätze

Asica-Corpus

← → www.asica.gwi.uni-muenchen.de/index.php?corpus=1 ☆ open access lmu



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

FAKULTÄT FÜR SPRACH- UND LITERATURWISSENSCHAFTEN
ROMANISCHE PHILOLOGIE
ATLANTE SINTATTICO DELLA CALABRIA





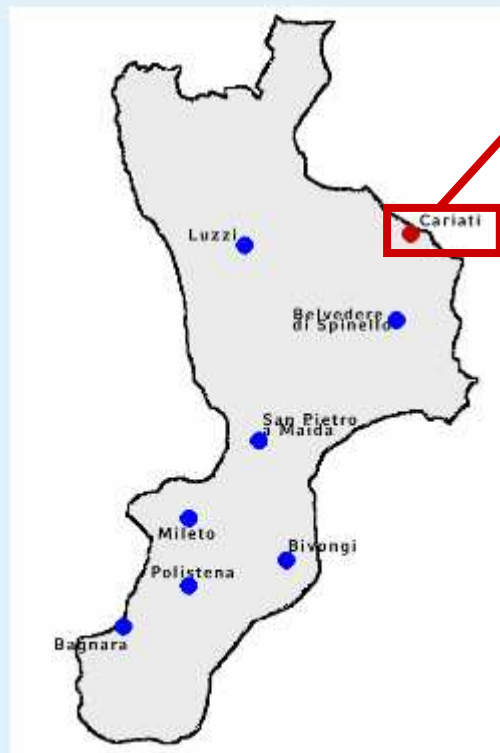
atlante
ricerca interattiva
campione
informanti
luoghi
corpus
testi
links
progetto
backend

Cliccare sulla sigla per consultare i materiali (trascrizione e soundfile)
Cliccare su 'CES' per avviare la presentazione secondo lo standard 'CES' (Corpus encoding standard)
 (per ascoltare installare il Plugin "QuickTime" o programma equivalente!)

Acc1mIQ1 (mp3 CES TG)	Bel1wID1 (mp3 CES TG)	Biv2wIQa (mp3 CES TG)	Luz2mIQa (mp3 CES TG)
Acc1mIQ2 (mp3 CES TG)	Bel1wII1 (mp3 CES TG)	Car1mDD1 (mp3 CES TG)	Luz2wDD1 (mp3 CES TG)
Acc1mIQa (mp3 CES TG)	Bel1wIQ1 (mp3 CES TG)	Car1mDDQ2 (mp3 CES TG)	Luz2wDI1 (mp3 CES TG)
Acc1wIQ1 (mp3 CES TG)	Bel1wIQ2 (mp3 CES TG)	Car1mDI1 (mp3 CES TG)	Luz2wDQ1 (mp3 CES TG)
Acc2mID1 (mp3 CES TG)	Bel1wIQa (mp3 CES TG)	Car1mDQ1 (mp3 CES TG)	Luz2wDQ2 (mp3 CES TG)
Acc2mII1 (mp3 CES TG)	Bel2mDD1 (mp3 CES TG)	Car1mDQa (mp3 CES TG)	Luz2wDQa (mp3 CES TG)
Acc2mIQ1 (mp3 CES TG)	Bel2mDI1 (mp3 CES TG)	Car1mID1 (mp3 CES TG)	Luz2wID1 (mp3 CES TG)
Acc2mIQ2 (mp3 CES TG)	Bel2mDQ1 (mp3 CES TG)	Car1mIQ1 (mp3 CES TG)	Luz2wII1 (mp3 CES TG)
Acc2mIQ3 (mp3 CES TG)	Bel2mDQ2 (mp3 CES TG)	Car1wDD1 (mp3 CES TG)	Luz2wIQ1 (mp3 CES TG)
Acc2mIQa (mp3 CES TG)	Bel2mDQa (mp3 CES TG)	Car1wDI1 (mp3 CES TG)	Luz2wIQ2 (mp3 CES TG)
Acc2wID1 (mp3 CES TG)	Bel2mID1 (mp3 CES TG)	Car1wDQ1 (mp3 CES TG)	Luz2wIQa (mp3 CES TG)
Acc2wII1 (mp3 CES TG)	Bel2mII1 (mp3 CES TG)	Car1wDQ2 (mp3 CES TG)	Mil1mDD1 (mp3 CES TG)
Acc2wIQ1 (mp3 CES TG)	Bel2mIQ1 (mp3 CES TG)	Car1wDQa (mp3 CES TG)	Mil1mDI1 (mp3 CES TG)
Acc2wIQ3 (mp3 CES TG)	Bel2mIQ2 (mp3 CES TG)	Car1wID1 (mp3 CES TG)	Mil1mDQ1 (mp3 CES TG)
Ama2wDD1 (mp3 CES TG)	Bel2mIQa (mp3 CES TG)	Car1wIQ1 (mp3 CES TG)	Mil1mDQ2 (mp3 CES TG)
Ama2wDI1 (mp3 CES TG)	Bel2wDD1 (mp3 CES TG)	Car1wIQ2 (mp3 CES TG)	Mil1mID1 (mp3 CES TG)
Ama2wDQ1 (mp3 CES TG)	Bel2wDI1 (mp3 CES TG)	Car2mDD1 (mp3 CES TG)	Mil1mII1 (mp3 CES TG)
Ama2wDQ2 (mp3 CES TG)	Bel2wDQ1 (mp3 CES TG)	Car2mDI1 (mp3 CES TG)	Mil1mIQ1 (mp3 CES TG)
Ama2wDQa (mp3 CES TG)	Bel2wDQ2 (mp3 CES TG)	Car2mDQ1 (mp3 CES TG)	Mil1mIQ2 (mp3 CES TG)
Bag1mDD1 (mp3 CES TG)	Bel2wDQ3 (mp3 CES TG)	Car2mDQ2 (mp3 CES TG)	Mil1mIQa (mp3 CES TG)
Bag1mDI1 (mp3 CES TG)	Bel2wDQa (mp3 CES TG)	Car2mDQa (mp3 CES TG)	Mil1wDD1 (mp3 CES TG)
Bag1mDI2 (mp3 CES TG)	Bel2wID1 (mp3 CES TG)	Car2mID1 (mp3 CES TG)	Mil1wDI1 (mp3 CES TG)
Bag1mDQ1 (mp3 CES TG)	Bel2wII1 (mp3 CES TG)	Car2mII1 (mp3 CES TG)	Mil1wDQ1 (mp3 CES TG)
Bag1mDQ2 (mp3 CES TG)	Bel2wIQ1 (mp3 CES TG)	Car2mIQ1 (mp3 CES TG)	Mil1wDQ2 (mp3 CES TG)
Bag1mID1 (mp3 CES TG)	Bel2wIQ2 (mp3 CES TG)	Car2mIQ2 (mp3 CES TG)	Mil1wID1 (mp3 CES TG)
Bag1mII1 (mp3 CES TG)	Bel2wIQa (mp3 CES TG)	Car2mIQa (mp3 CES TG)	Mil1wII1 (mp3 CES TG)
Bag1mIQ1 (mp3 CES TG)	Biv1mDD1 (mp3 CES TG)	Car2wDD1 (mp3 CES TG)	Mil1wIQ1 (mp3 CES TG)
Bag1mIQ2 (mp3 CES TG)	Biv1mDQ1 (mp3 CES TG)	Car2wDI1 (mp3 CES TG)	Mil1wIQ2 (mp3 CES TG)
Bag1mIQa (mp3 CES TG)	Biv1mDQ2 (mp3 CES TG)	Car2wDQ1 (mp3 CES TG)	Mil1wIQa (mp3 CES TG)
Bag1wDD1 (mp3 CES TG)	Biv1mDQa (mp3 CES TG)	Car2wDQ2 (mp3 CES TG)	Mil2mDD1 (mp3 CES TG)
Bag1wDI1 (mp3 CES TG)	Biv1mID1 (mp3 CES TG)	Car2wDQa (mp3 CES TG)	Mil2mDI1 (mp3 CES TG)
Bag1wDI2 (mp3 CES TG)	Biv1mII1 (mp3 CES TG)	Car2wID1 (mp3 CES TG)	Mil2mDQ1 (mp3 CES TG)
Bag1wDI3 (mp3 CES TG)	Biv1mIQ1 (mp3 CES TG)	Car2wII1 (mp3 CES TG)	Mil2mDQ2 (mp3 CES TG)

Einzelwort- und Belegrecherche im Korpus (AsiCa)

Ricerca interattiva



Lemma

p

ok

reset

P.ESCLAMATIVO ()

padre (NPar)

paese (N)

pagare (V)

paio (N)

palla (N)

pallone (N)

pane (N)

papa (NPar)

paradiso (N)

parcella (N)

parecchio (DI)

parente (NPar)

Treffer 1 von 48



Gehe zu Treffer:

Wort: pan - **ID:** 18302 - **Interview:** Car1mDQ1 - **Text-Nr:** 12 - **Sprecher:** I:

Mentre correva è caduto . | . | menṭri ka fujia sugnu kadut | . | Vieni a prenderti il pane! |

vena pigghjarata u [pan](#)

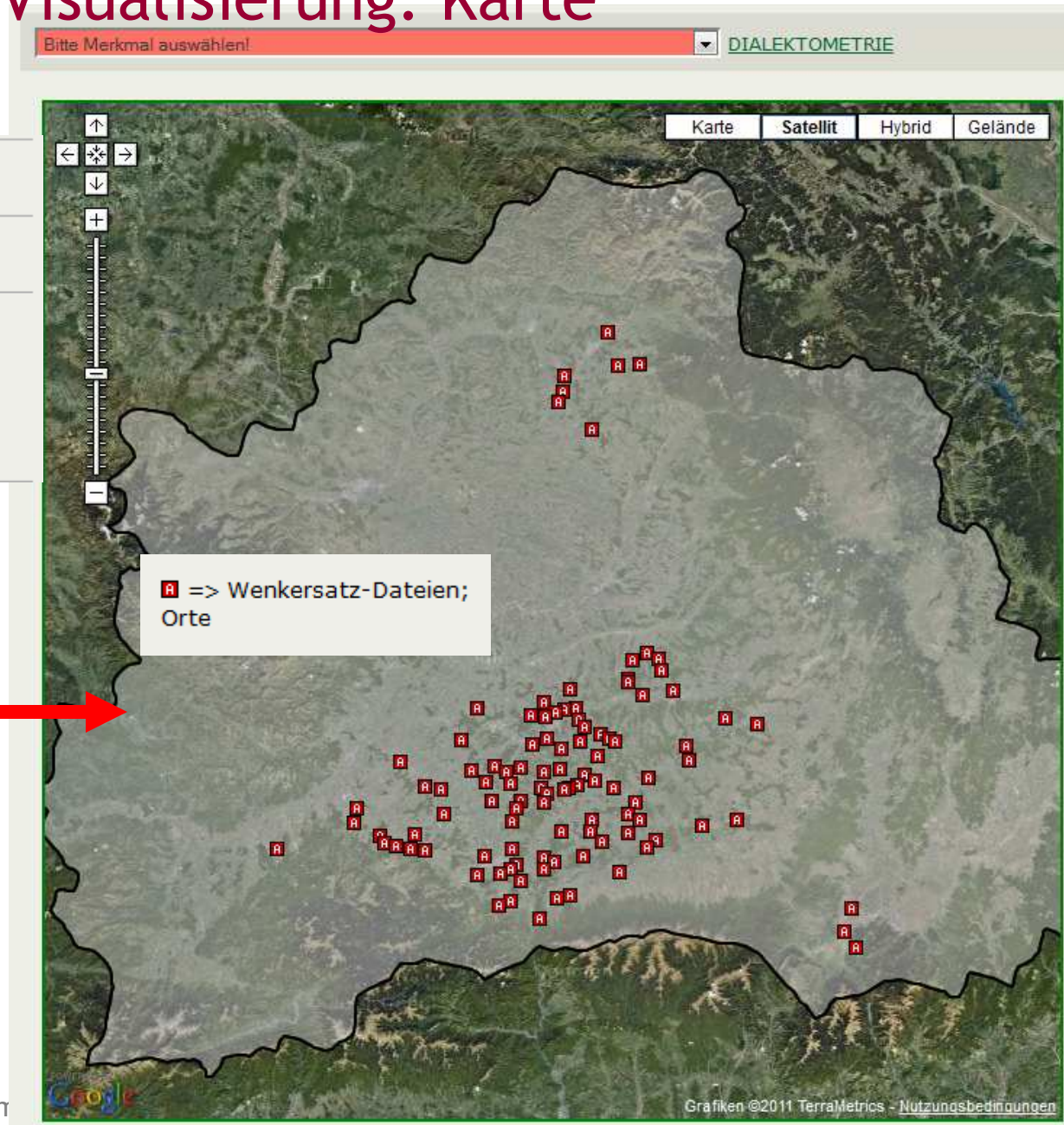
Per lavarsi è dovuto uscire fuori . | pe mi llavara [h]addi niḡḡira fore | per lavarsi | aah | . |

Visualisierung: Karte

ÜBER DAS PROJEKT
CORPUS
WENKERSATZANALYSE
Liste
Karte

Google Maps

ASD-Datenbank



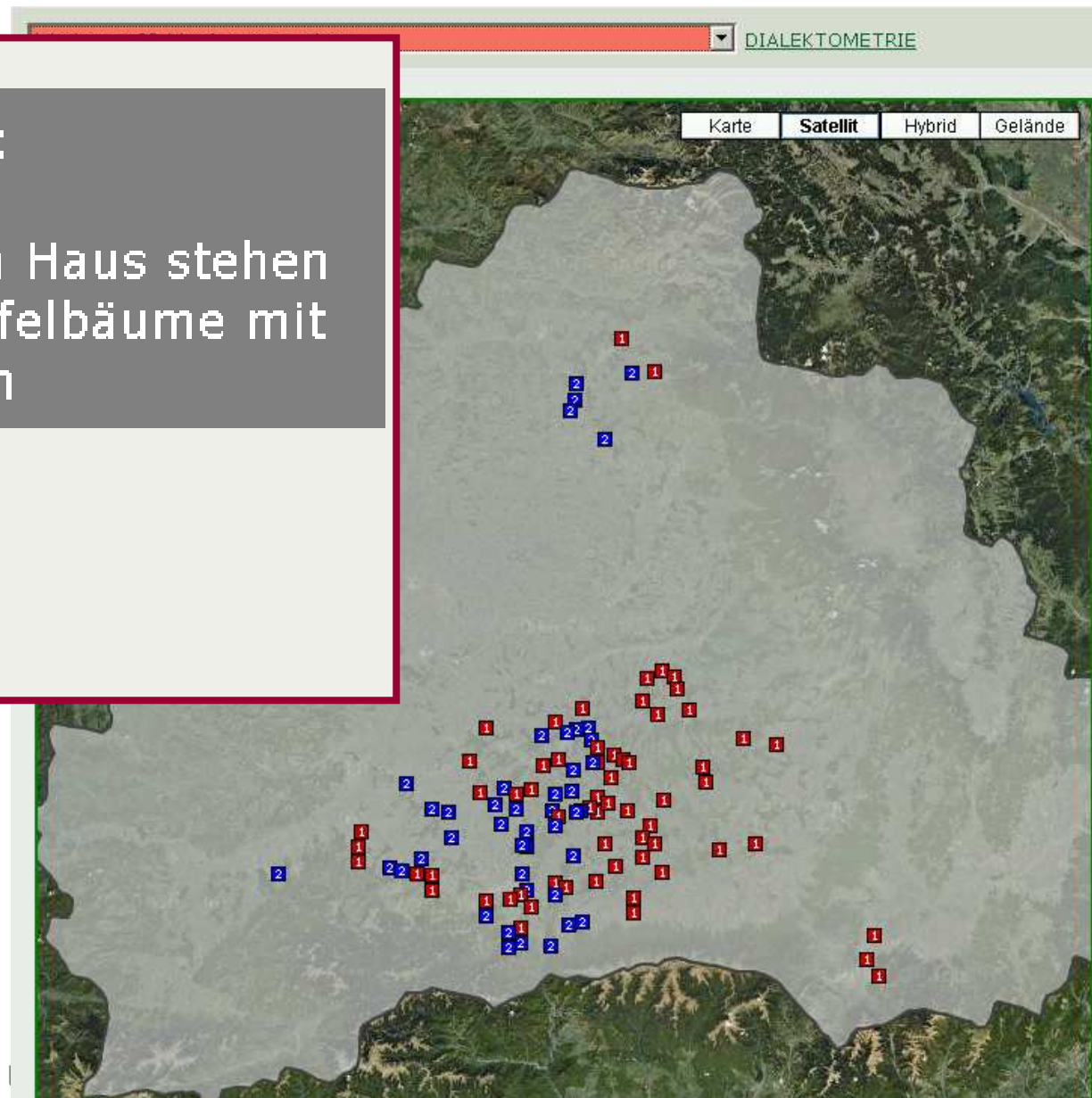
Beispiel qualitative Karte

Wenkersatz 26:

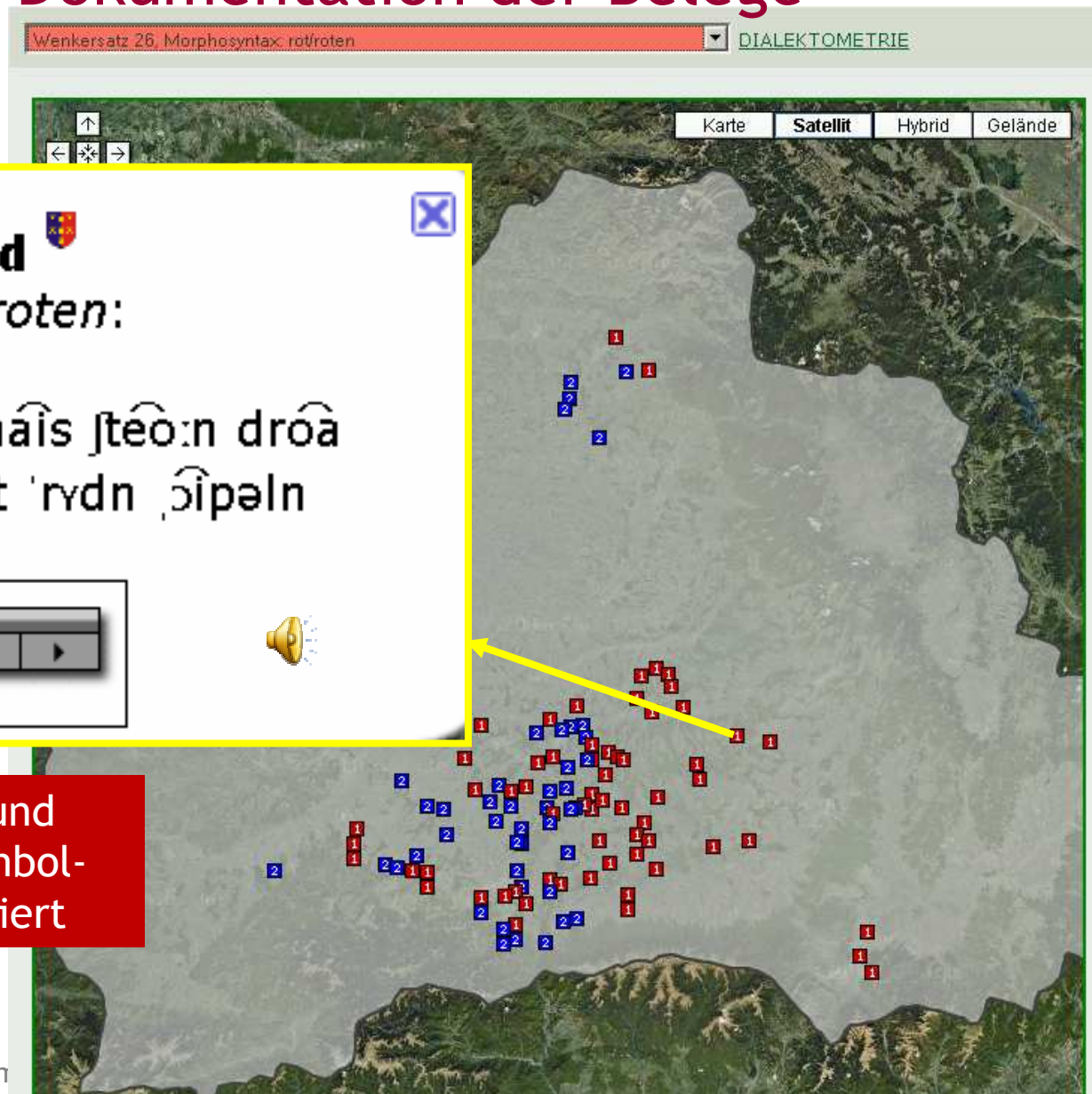
Hinter unserem Haus stehen
drei schöne Apfelbäume mit
roten Äpfelchen

1 => roten

2 => rot



Beispiel qualitative Karte: Dokumentation der Belege



Vorteile analytischer und
synthetischer Punkt-Symbol-
Karten werden kombiniert

Qualitative Karte mit arealtypologischer Auswertung der Belege

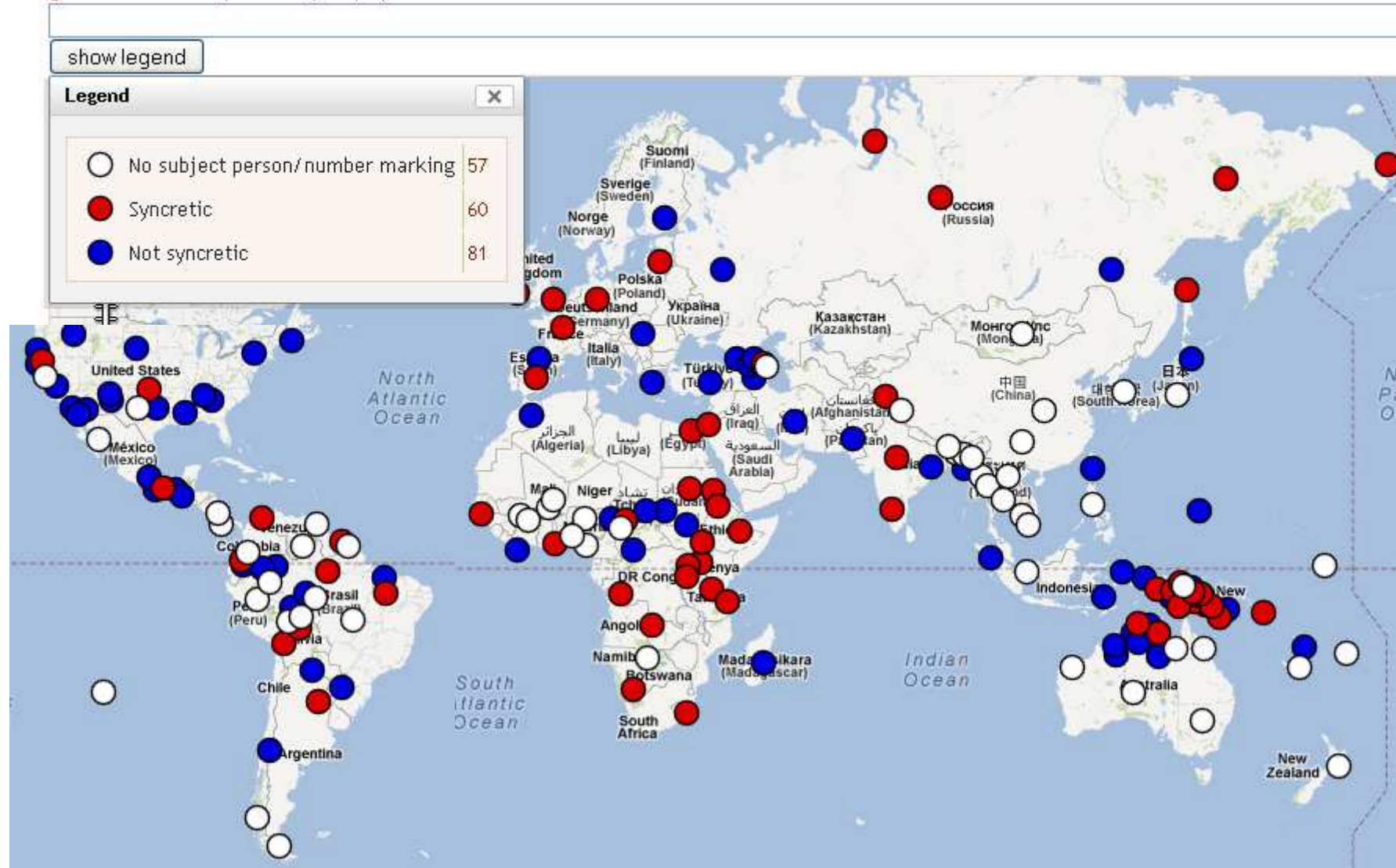
- Synthetische Kartierung komplexer Informationen
- Beispiel AdIS
 - Präsensparadigma von it. *essere* ‚sein‘

The World Atlas of language structures (<http://wals.info>)

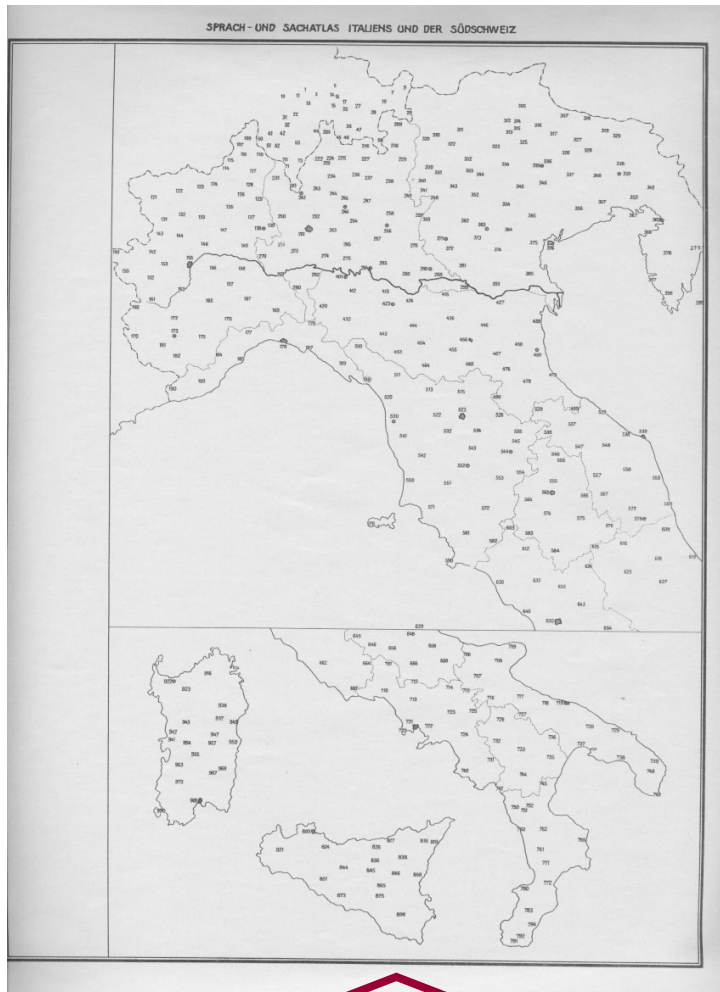
Feature 29A: Syncretism in Verbal Person/Number Marking

by Matthew Baerman and Dunstan Brown

get URL for the map currently displayed



Datenbasis AIS 1689 a



in dieser Weise
für alle 412 Atlaspunkte

Legende:

Tabelle α enthält die Präsensformen, Tabelle β die Imperfektformen; Tabelle γ die Konj. Präsens Formen.

α) Präsens von essere (presente di essere)

Hinsichtlich der Art der Aufnahme der hier gebotenen Präsensformen sehe man die Legende der K. VIII, 1682 ein. Die Trennung der pronominalen Elemente machte hier wenig Schwierigkeiten. Die Subjektspronomina sind auf K. VIII, 1683 einzusehen. Es ist unmöglich, hier alle Karten, auf denen Präsensformen von 'essere' figurieren, aufzuzählen. Man konsultiere den Generalindex des AIS. Wir begnügen uns hier, für jede Person des Präsens von 'essere' auf eine Karte hinzuweisen.

<u>sono</u>	: K. IV, 700	148,6 = 94,14 = 34,13
<u>sei</u>	: K. IV, 712	156,8 = 98,3 = 0
<u>è</u>	: K. VII, 1450	55,3 = 38,2 = 11,8
<u>siamo</u>	: K. VIII, 1646	161,6 = 101,2 = 0
<u>siete</u>	: K. VIII, 1598	47,11 = 33,6 = 0
<u>sono</u>	: K. VIII, 1555	125,7 = 82,9 = 0

1 sùn; òys; òy
3 sùn; èys; èy
9 sùn; ès; ès
10 sùn; àys; è
11 sùn; àys; ày
13 sùn; àys; ày
14 sùn; ès; è

èsan; èsas; òyn
èsan; èsas; èyn
èsan; èsat; sùn
èsan; èsas; èn
èsan; èsas; àyn; èn
èsan; èsas; àyn
èsan; èsas; èn

[...]

Entwicklung des Paradigmas gegenüber dem Klassischen Latein

40

- unregelmäßiges Paradigma von lat. *esse* 'sein'
personaldeiktisch

	Origo		
Singular einsilbig	1. Person lat. <i>sum</i>	2. Person lat. <i>es</i>	3. Person lat. <i>est</i>
Plural zwei-/einsilbig	4. Person lat. <i>sumus</i>	5. Person lat. <i>estis</i>	6. Person lat. <i>sunt</i>
Singular einsilbig	1. Person it. <i>sono</i>	2. Person it. <i>sei</i>	3. Person it. <i>è</i>
Plural zweisilbig	4. Person it. <i>siamo</i>	5. Person it. <i>siete</i>	6. Person it. <i>sono</i>



Faktoren der Paradigmenentwicklung

- Synkretismen
 - Generalisierung eines der beiden Stämme und Verdrängung des anderen
 - Einsilbig- und Zweisilbigkeit der Formen
- Interpretation, z.B. im Sinne der Natürlichkeitstheorie

Verallgemeinerung der Zweisilbigkeit

Sonderfälle [1]



The image shows a map of Southern Italy, specifically the Gargano Peninsula. A callout box points to the location of San Giovanni Rotondo. The map includes labels for 'Slovenija', 'Pécs', 'Adriatic Sea', 'Foggia', 'Andria', 'Bari', and 'Shq'. Various colored squares (A, B, C, K) are marked on the map, likely indicating different linguistic or dialectal regions.

informante n. 708
carta n. 1689_2
comune/luogho: San Giovanni
Rotondo

'sonna; 'sinna; 'jenna; 'si:ma; 'si:ta;
'sonna
[so(/nne1; si(/nne1; ye(/nne1;
si-/me1; si-/te1; so(/nne1]

Kombination der synkretistischen Typen A und C (am Beispiel des Dialekts von Ravenna, AIS P 232)⁴³

Sonderfälle [2]

		personaldeiktisch		
		Origo		
Singular	1. Person	$\int \mathfrak{d}$	2. Person	3. Person
			$\int \tilde{i}$	$\mathfrak{E}$
Plural	4. Person	$\int \tilde{\epsilon}$	5. Person	6. Person
			$\int \tilde{i}$	$\mathfrak{E}$
		Typ C		Typ A

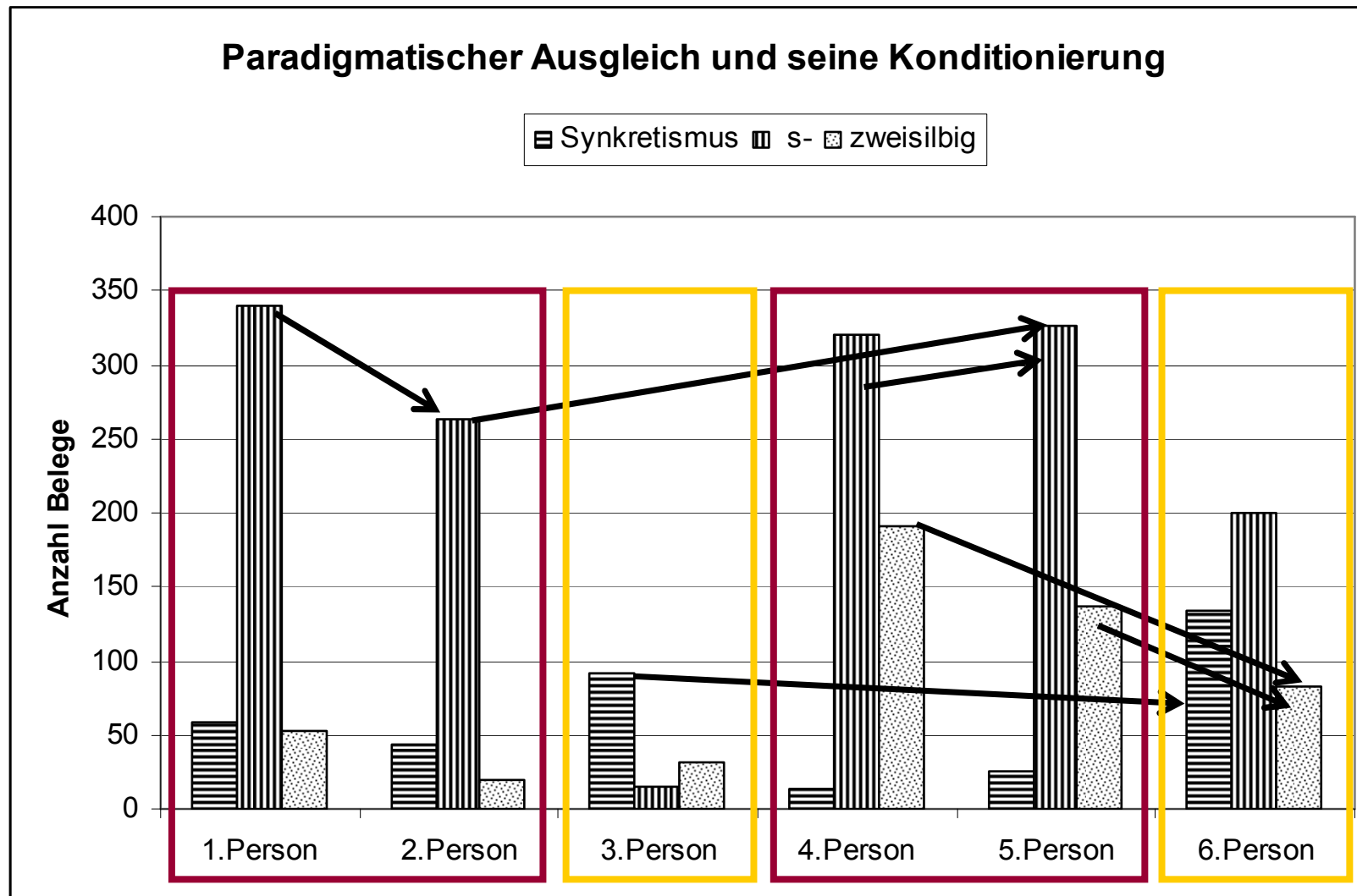


1., 2., 3., 4., 5., 6. Person

A	=>	001001	i	(77)
B	=>	100001	i	(47)
C	=>	010010	i	(23)
D	=>	011000	i	(8)
E	=>	100100	i	(6)
F	=>	011001	i	(5)
G	=>	110000	i	(3)
H	=>	000110	i	(2)
I	=>	000101	i	(2)
J	=>	010100	i	(2)
K	=>	111000	i	(1)
L	=>	001101	i	(1)
M	=>	100101	i	(1)
N	=>	010001	i	(1)

'1' steht für gemeinsame Form
hier: 2. = 6. Person



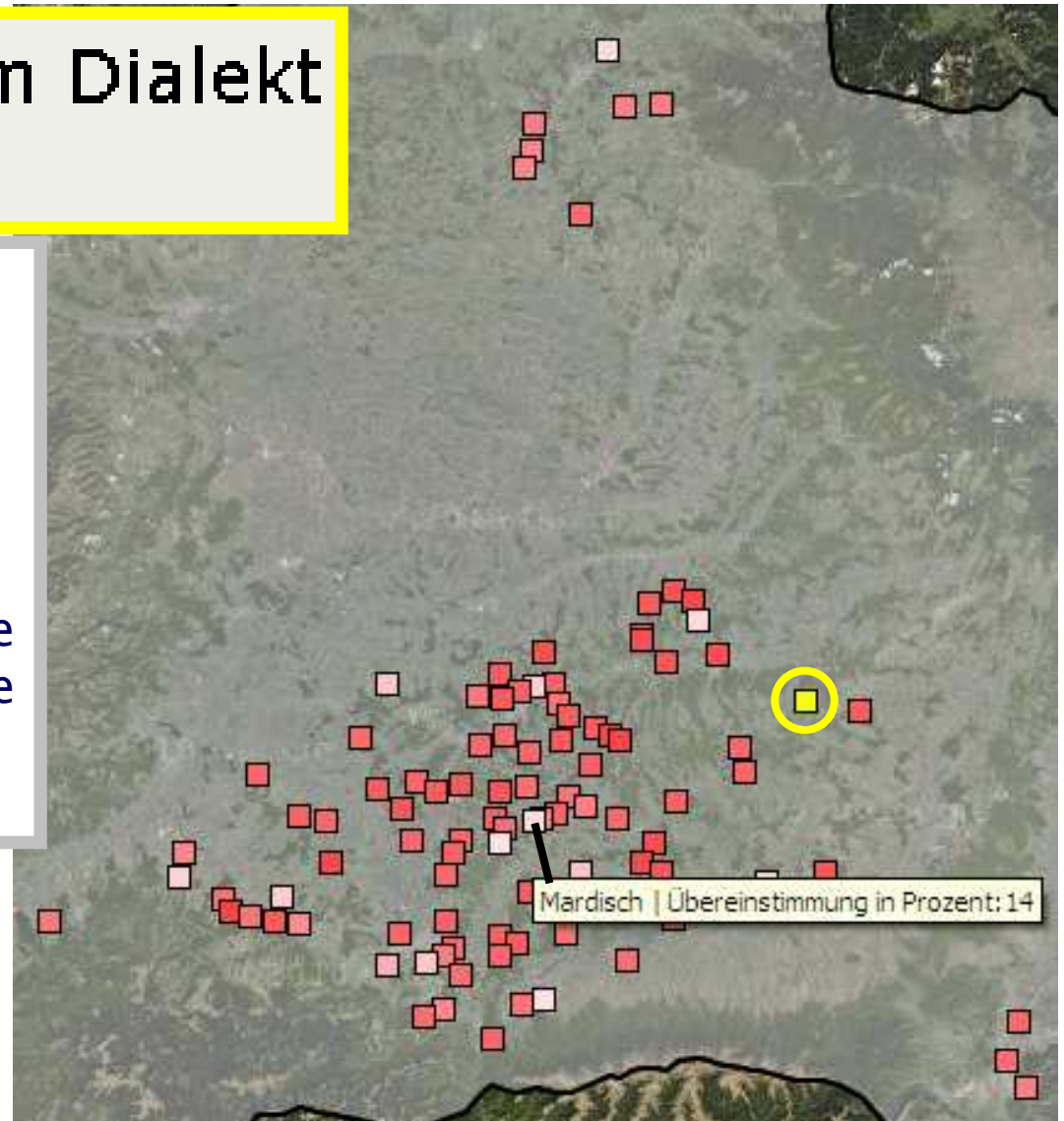


Beispiel quantitative Karte (Dialektometrie)

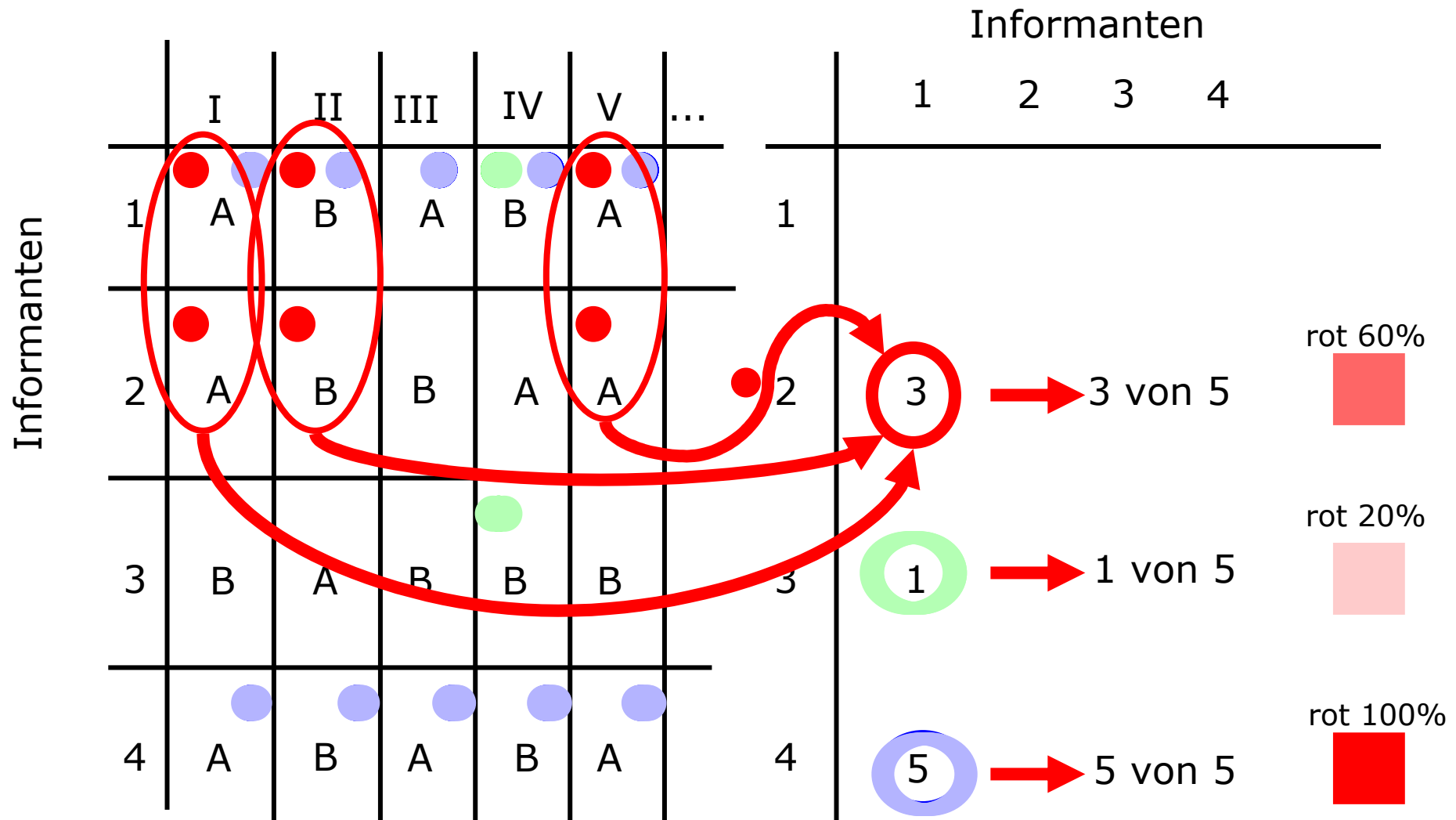
Ähnlichkeiten mit dem Dialekt von Keisd

je ähnlicher die Farbe,
desto mehr Merkmale
hat ein Ort mit dem
Bezugspunkt gemeinsam

noch nicht sehr viele Merkmale
erfasst, dennoch auffällig geringe
Arealbildung (z.B. keine sehr klare
Korrelation zwischen Ähnlichkeit
und Entfernung)



Quantitative Analyse



synoptische Kartierung: qualitativ vs. quantitativ

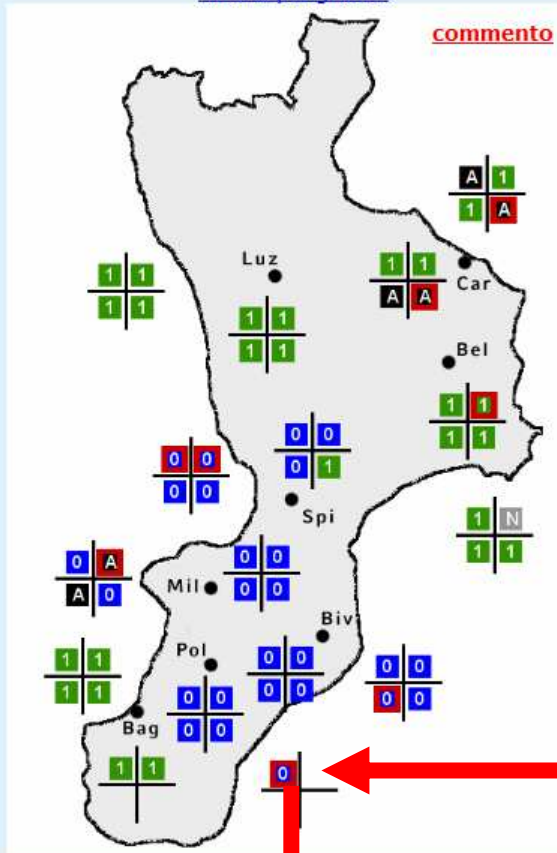
complemento verbale di 'volere'

F22: Un giorno vorrei ritornare al mio paese. -> ottativa, infinito dopo volere; posizione del possessivo; esistenza del condizionale

Cliccare sui simboli per vedere le risposte degli informanti!

Cartina basata sui questionari

[cartina più grande](#)

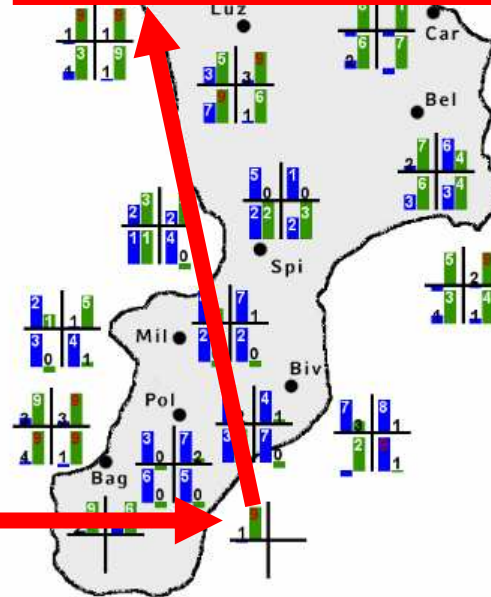


[Vedere tutte le risposte degli informanti](#)

Cartina basata sul materiale spontaneo

[cartina più grande](#)

Pol1mDD1	477	voggrju	andare	nenti pò bbasta mmo m'ndi voggrju
Pol1mDD1	450	vogliu	andare	mmo vogliu andare via
Pol1mDD1	167	voli	sapiri	eh ki aṭr ki voli sapiri
Pol1mDD1	168	voli	sapiri	voli sapiri na kosa de lla mia vita\$?
Pol1mDD1	446	voli	sapiri	e ppo bbasta non vi kuntu nenti per'k



Una '9' rossa significa più di nove risultati

[Vedere solo la cartina basata sui questionari](#)

- 0 senza infinito (La categoria)
- 1 con infinito
- A ambiguo
- N nessun risultato
- Variazione individuale condizionata dalla tecnica di elicitazione (questionario)

	interview	intervall	parola 1	parola 2	contesto
Pol1mDQ1	164	vorria	tuarnu	nu juarnu vorria u tuarnu ô paisi miu	

Multimedialität - Beispiel ASD

Datenpräsentation:

- akustisch
- doppelte Textgestalt (phonetische und/oder orthographische Transkription)
- qualitative und quantitative Kartierung: Verbreitung von Einzelmerkmalen und Gemeinsamkeiten lokaler Varietäten durch gemeinsame Merkmale
- Option: Aufbereitung mit Einzelwort- bzw. Einzelformrecherche (→ Tagging; derzeit 81.818 Tokens, davon 9.696 insgesamt 2.721 Types zugeordnet, die 339 Lemmata zugeordnet sind)
 - deshalb auch lexikographisch interessant

Internetoberfläche des ASD

Internetoberfläche des ASD (Screenshot)

Browser: Mozilla Firefox
URL: <http://siebenbuergen.gwi.uni-muenchen.de/index.php?karte=true>

Logo: LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Institut: INSTITUT FÜR DEUTSCHE KULTUR UND GESCHICHTE SÜDOSTEUROPAS
INSTITUT FÜR ROMANISCHE PHILOGIE
IT-GRUPPE GEISTESWISSENSCHAFTEN

ASD AUDIO-ATLAS SIEBENBÜRGISCH-SÄCHSISCHER DIALEKTE

LMU-Startseite Login

ÜBER DAS PROJEKT
CORPUS
WENKERSATZANALYSE
Liste
Karte
DOKUMENTE

Wenkersatz 14:
Mein liebes Kind bleib hier
unten stehen die bösen
Gänse beißen dich sonst tot

1 => liebes/liebet
2 => lieb

Wenkersatz 14, Morphosyntax: liebes/lieb (flektiert – nicht flektiert) DIALEKTOMETRIE

Karte Satellit Hybrid Gelände

Map showing distribution of dialect forms (1 and 2) across the region.



Corpus - Beispiel Wenkersatztranskription

LMU-Startseite asduser [logout](#)

CORPUS

ÜBER DAS PROJEKT
WENKERSATZANALYSE
DOKUMENTE

Anhören mit Flashplayer
Anhören mit Quicktime-Player

asduser
h3rmann5tadt

Datenbestand: 2212 | Anzahl gefundene Dateien: 1
[Alle Dateien](#) | [Nur transkribierte Dateien](#) | [Alle Filter entfernen](#) | [Karte bei Google Earth](#)

Audio	Ort	Jahr	Alter	Corpus	Stichwörter
1113-05	Filter	Filter	Filter	Filter	Filter
	Abtsdorf bei Agnetheln (S)	1972	32	Bukarest	Wenkersätze

Phonetische Transkription
[Anmerkungen](#) | [Textgrid-Datei](#)

[1]: ɪç hɪsə ʃʊstər johana bæ fun ɔpəsdrɛf
 [2]: fɔ vənəɪfɔmhaər
 [3]: vɛŋkərztʃɛ
 [4]: {WS Anfang}
 [5]: {WS 1} uēht
 [6]: ## æm 'væŋtər 'flejən də gə'drɛçt 'bladər æn
 dər 'lɔft ə rəm
 [7]: {WS 2} tʃvæ:
 [8]: ## ət hɪɪrt glɛç 'oäf fum 'ʃno:ən drɛ vɪd dət
 'vadəɪr 'vedər 'beɪsər
 [9]: {WS 3} dro:
 [10]: ## deä 'kiʊln æn 'iūvən dət 'ə 'mæltʃ 'bɔ:ld
 'ufɛɪt tʃə 'kəðxən
 [11]: {WS 4} fər
 [12]: ## dər 'geät 'ɔ:ld 'mæʊn æs mæɪt 'əm 'rəʊs
 dərç 't 'aɪs gə brəðxən ən æs æn 't 'kəʊlt 'vɔsər
 gə fɔlən
 [13]: {WS 5} fɔf
 [14]: ## ə æs fɪr 'fər ɔvər 'ze:s 'vəðxə gə ʃtɔrvən
 [15]: {WS 6} ze:ɪs
 [16]: ## dət 'faiər vəʊr tʃə 'ʃtɔrk dər 'ke də
 'keəðxə zən ə o 'anʃdə 'ʃvɔrtʃ fər brɔ:t
 [17]: {WS 7} zɪvən

Orthographische Transkription
[Anmerkungen](#) | [Textgrid-Datei](#)

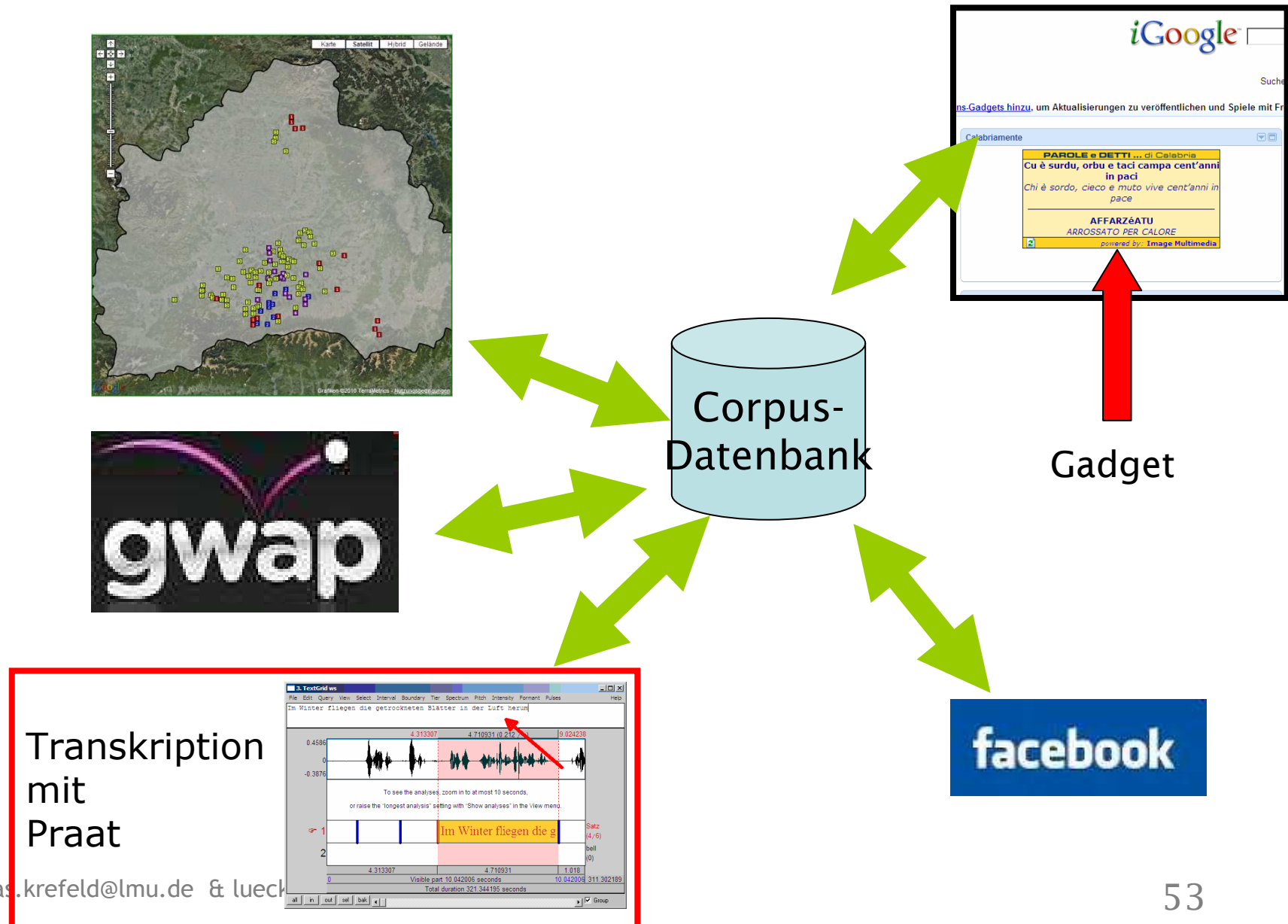
[Wenkersätze](#)

1. Im Winter fliegen die trockenen Blätter in der Luft herum.
2. Es hört gleich auf zu schneien, dann wird das Wetter wieder besser.
3. Tu Kohlen in den Ofen, dass die Milch anfängt zu kochen.
4. Der gute alte Mann ist mit dem Pferd durchs Eis gebrochen und in das kalte Wasser gefallen.
5. Er ist vor vier oder sechs Wochen gestorben.
6. Das Feuer ist zu stark, der Kuchen ist ja unten ganz schwarz verbrannt.
7. Er isst die Eier immer ohne Salz und Pfeffer.
8. Die Füße tun mir sehr weh, ich glaube, ich habe sie mir aufgerieben.
9. Ich bin bei der Frau gewesen und habe es ihr gesagt, und sie sagte, sie will es auch ihrer Tochter sagen.
10. Ich will es nie mehr tun.
11. Ich gebe dir gleich eins mit dem Kochlöffel über die Ohren, du Affe.
12. Wo gehst du hin? Sollen wir mit dir kommen?
13. Es waren schlechte Zeiten.
14. Mein liebes Kind, bleib hier unten stehen, die

Inhalt

1. Ein Blick in die Wissenschaftsgeschichte vor dem Hintergrund der Neuen Medien
 2. Unsere Projekte
 3. Aspekte der Digitalisierung
 4. Daten und Metadaten
 5. Analyse und Visualisierung
 6. Web 2.0 - Kollaboratives Arbeiten
-

Web 2.0: Interaktion & Zusammenarbeit





Observatorium aktueller Entwicklungen des
Regionalitalienischen

<http://beta.metropolitalia.org/>

Konzept: Play4Science www.play4science.org/

Entwicklungsphase, Start im Juli 2012

Was wird von Sprechern des Italienischen für
typisch für welche Region / welchen Ort und
welchen Sprechertyp gehalten gehalten?

COME GIOCARE AREA COMMENTI REGISTRATI ACCEDI

metropol
italia **BETA**

Mi ha dato una buffettuna! ?

L'espressione significa: Mi ha dato uno schiaffo!

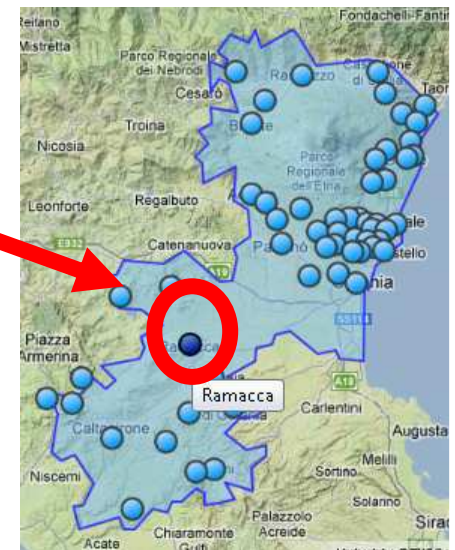
X Non lo so. Passo

Turno 3 su 3

Punti
Viene da: 0

Clicca sulla cartina oppure
inserisci la tua risposta qui:





Turno 3 su 3

Punti 0

Scommessa: Quale sarà la percentuale dei giocatori che conosce l'origine della espressione?

Turno 3 su 3

Punti 3

Complimenti, 46% dei giocatori la pensa come te, hai ottenuto 3 punti.



1

Cosa te lo fa pensare? Clicca sugli elementi che hanno influenzato la tua scelta.

Turno 3 su 3

Punti 13

Complimenti, 100% dei giocatori la pensa come te.



3

Chi userebbe questa espressione? Se rispondi, potresti ottenere dei punti bonus.

Mi ha dato una buffettuna! L'espressione significa: Mi ha dato uno schiaffo!

Sesso
 Età
 Livello di istruzione

Turno 3

Mi ha dato una buffettuna! 16 Punti

La mia scommessa: 90% pensa sia tipica di Catania

La mia scelta: Catania

Confermazione: 60% dagli tutti

Cosa puntano tutti:
 1 su 4 (25%) pensa sia tipica di: Torino.
 1 su 4 (25%) pensa sia tipica di: Sud.
 1 su 4 (25%) pensa sia tipica di: Catania.
 1 su 4 (25%) pensa sia tipica di: Calabria.

3 Punti 13 Punti



1

Cosa te lo fa pensare? Clicca sugli elementi che hanno influenzato la tua scelta.

Mi ha dato una buffettuna!

✓ OK

2

Turno 3 su 3

Punti

13

Complimenti, 100% dei giocatori la pensa come te.

3

Chi userebbe questa espressione? Se rispondi, potresti ottenere dei punti bonus.

X Salta questo passaggio

Mi ha dato una buffettuna! ?

L'espressione significa: Mi ha dato uno schiaffo!

Sesso

femminile maschile

Età

giovane anziano

Livello di istruzione

poco istruito istruito

✓ Avanti



Turno 3

Mi ha dato una buffettuna! ?

16
Punti

La mia scommessa:

90% pensa sia tipica di
Catania

Confermazione:

60% dagli tutti

3 Punti

La mia scelta:

Catania

Cosa puntano tutti:

1 su 4 (25%) pensa sia tipica di: Torino.
1 su 4 (25%) pensa sia tipica di: Sud.
1 su 4 (25%) pensa sia tipica di:
Catania.
1 su 4 (25%) pensa sia tipica di:
Calabria.

13
Punti

Gioca di nuovo

Aggiungi una nuova espressione

Scommessa: Quale sarà la percentuale dei giocatori che conosce l'origine della tua espressione?

 10

Si dice in:

Clicca sulla cartina oppure inserisci la tua risposta qui:

Aggiungi un'espressione che ritieni tipica di un luogo o di una determinata regione.

L'espressione significa:

✓ **Invia la tua espressione**



Spiel „Accenti Urbani“

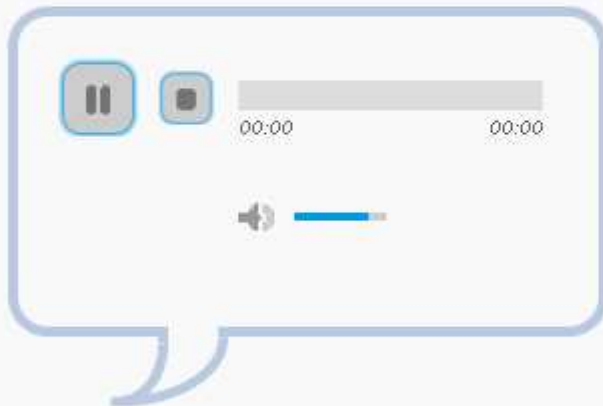
accenti urbani

Clicca sulla cartina per scegliere una città.

Turno 1 su 10

Punti

0



Da che città vengono questi parlanti?

INFORMAZIONI VAI ALLA CLASSIFICA



Anleitungen verfügbar

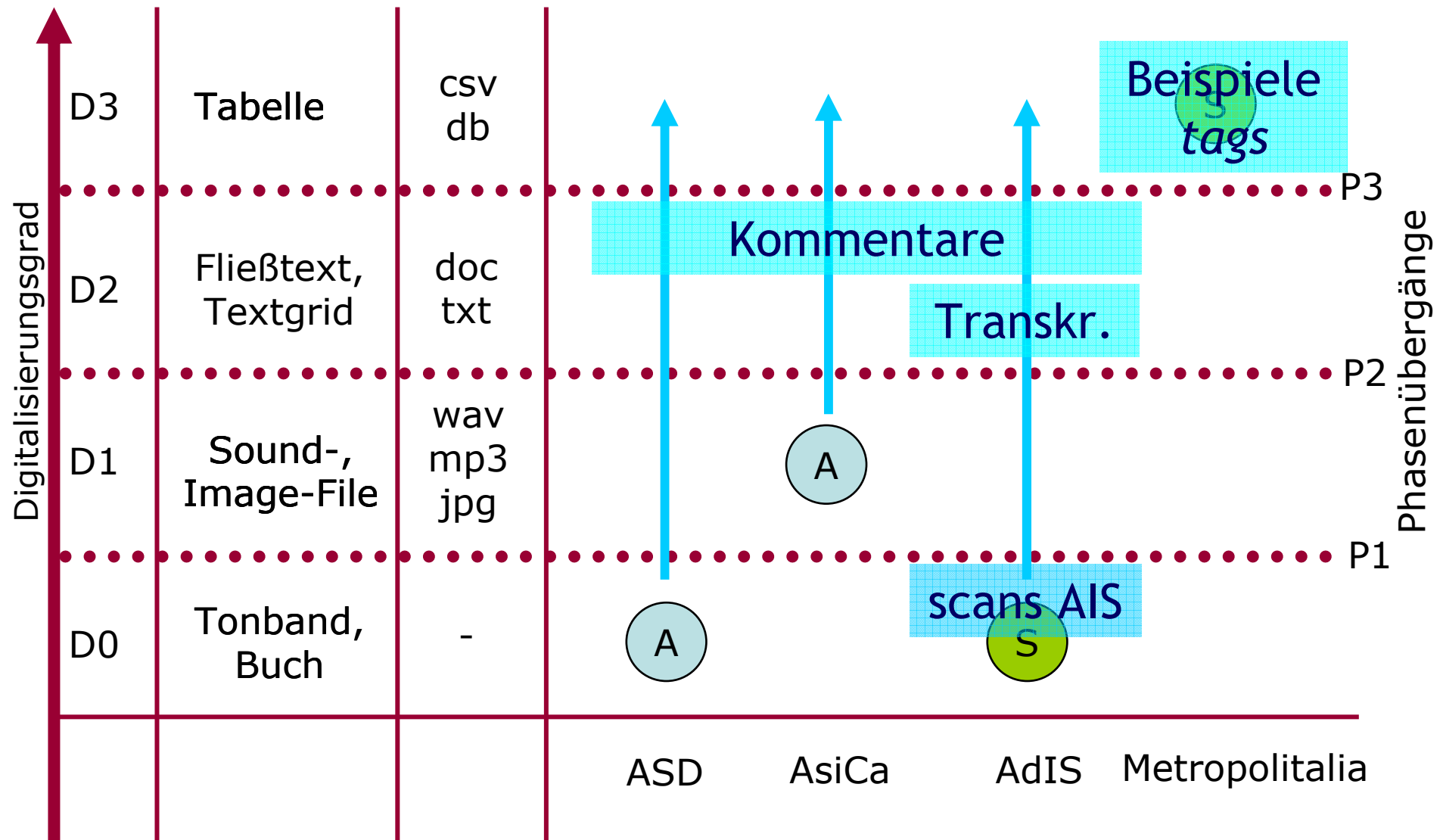
➤ Hinweise Datenerfassung

E20					
A	B	C	D	E	F
	Beschreibung Merkmal	Beschreibung Variante	abfrage	Kommentar Merkmal	Kommentar Varianten
MUSTER	carta 1079 (capra # fonetica)	k-	select id_comune, id_informanti_ais, ais_nr, ais_index, comuni.name as ort, aeusserung as kriterium from comuni join informanti_ais using (id_comune) join aeusserungen_ais using (id_informanti_ais) where sachaufnahmen != 1 and ais_karte = 1079 and (aeusserung like '% k%' or aeusserung like 'k%')	Kommentar zum Merkmal nur in die jeweils erste Zeile schreiben!	Kommentar zu Variante 1
	carta 1079 (capra # fonetica)	tsch-	select id_comune, id_informanti_ais, ais_nr, ais_index, comuni.name as ort, aeusserung as kriterium from comuni join informanti_ais using (id_comune) join aeusserungen_ais using (id_informanti_ais) where sachaufnahmen != 1 and ais_karte = 1079 and (aeusserung like '% c/%' or aeusserung like 'c/%')		Kommentar zu Variante 2
	carta 1079 (capra # fonetica)	tj-	select id_comune, id_informanti_ais, ais_nr, ais_index, comuni.name as ort, aeusserung as kriterium from comuni join informanti_ais using (id_comune) join aeusserungen_ais using (id_informanti_ais) where sachaufnahmen != 1 and ais_karte = 1079 and (aeusserung like '% c)%' or aeusserung like 'c)%')		Kommentar zu Variante 3
	carta 1079 (capra # fonetica)	ts-	select id_comune, id_informanti_ais, ais_nr, ais_index, comuni.name as ort, aeusserung as kriterium from comuni join informanti_ais using (id_comune) join aeusserungen_ais using (id_informanti_ais) where sachaufnahmen != 1 and ais_karte = 1079 and (aeusserung like '% ts%' or aeusserung like 'ts%')		Kommentar zu Variante 4

➤ Richtlinien Kommentare

➤ Transkriptionsregeln usw.

Kollaborative Optionen ...



© thomas.krefeld@lmu.de & luecke@lmu.de