Torsten Andreas
Ingo Fehrmann
Nicole Schumacher
Julia Welchering

Sprachproduktion japanischer DaF-Lernender im Unterricht

Personen
> 40 Studierende der Tokai-Universität Tokio im A1-DaF-Kurs

Metadaten: Alter, Geschlecht, Studienfach
Sprachbiographie, C-Test-Ergebnis

Sprachdaten: Fragebogen zur Sprachbiographie, Prä- und Posttests zu Syntax (Satzgliedpuzzle), lokalen Relationen (Lückentext zur Bildbeschreibung) sowie C-Tests
(Fragebögen und Tests noch nicht digitalisiert)

4 Unterrichtsgruppen

Metadaten: Sprachbiographie der Lehrenden
Kursinhalte/ Curriculum, verwendete Lehrwerke

Sprachdaten: Lehrersprache (Video- und Audiodaten, s.u.)
Unterrichtsmaterialien inkl. visuelle Elizitierungsstimuli
Aufgabenstellungen (Publikation in Vorbereitung)
z.T. Tafelanschriebe (noch nicht digitalisiert)

Pilotstudie:
20 Studierende der Tokai-Universität Tokio [1]

Vergleichsgruppe:
20 Studierende der HU Berlin mit L1 Deutsch – Aufgabe Skype-Termin

METHODEN

keine Kontrolle / Manipulation	mittlere Kontrolle / Manipulation	starke Kontrolle / Manipulation
Beobachtung in „authentischen“ Situationen	Elizitierung durch fokussierte, bedeutungszentrierte Aufgaben	Elizitierung durch formenzentrierte Aufgaben oder Experimente
„natürliche“ Daten	elizitierte Daten	experimentelle Daten

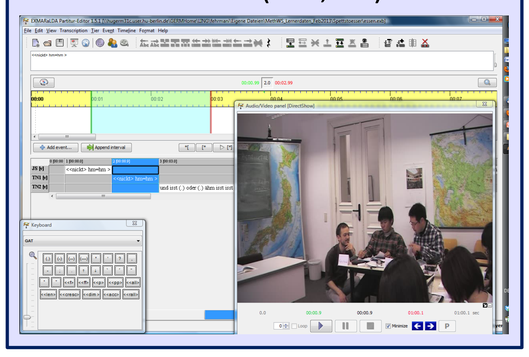
Aufgaben:
- Bericht
- Skype-Termin
- To-Do-Liste
- Bildergeschichte
- Krimigeschichte
- Esskultur
- Bildbeschreibung
- Ostereier verstecken
- Wettervergleiche

„Natürlichkeit“[2]
„Kontextualisierung“[3]

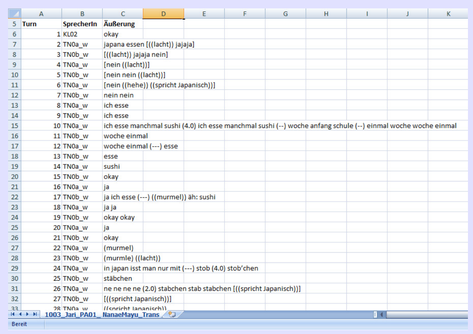
ARTEN VON LERNERDATEN

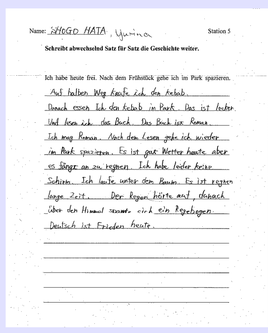
Lernersprachliche Daten

„natürliche“ Sprachdaten:
Videodaten
Audiodaten (mp3, z.T. wav)
GAT-Transkripte (exb, xls)



Elizitierte Sprachdaten:
Videodaten
Audiodaten (mp3, z.T. wav)
GAT-Transkripte (exb, xls)
handschriftliche Texte (noch nicht digitalisiert)
Zeichnungen (nach Bildbeschreibung)





Geplanter Zugang

Medienrepositorium der HU:

- Privat (derzeit): Arbeitsdateien vor Anonymisierung
- Öffentlich (geplant):
  Nach kompletter Anonymisierung
  - Audiodateien
  - Videodateien
  - GAT-Transkripte

JaDe-Korpus (in Progress):

- Nach Aufbereitung erster, ausgewählter Daten öffentlich, Zugriff über ANNIS (browserbasiert)
- (vorläufige) Ebenenstruktur:
  - orthographisch diplomatische Transkription
  - orthographisch normalisierte Transkription (deutsch/ japanisch)
  - Pausen
  - extralinguistische Ereignisse
  - Wortartenannotation nach erweitertem STTS (TreeTagger, manuell)
  - Lemmatisierung
  - *Disfluencies*: Unterbrechungen im Redefluss
  - *Repair*: (Selbst-)Reparaturen
  - *Backchanneling*: verbale/ nonverbale Kooperationsrückversicherungen