MULTILIT 2010-2013

Development of Oral and Written Abilities in L1, L2 and L3 by Multilingual Children and Adolescents with Turkish Background in France and Germany

Overview

1. Project partner
2. Empirical Research Questions
3. Methodology
     3.1 Project Design
     3.2 Data collection
     3.3 MULTILIT Germany
     3.4 Analysis Criteria
     3.5 Automatic POS-Tagging
     3.6 MULTILIT corpus

1. Project partner

A cooperative project

France: Laboratoire Dynamique Du Langage (UMR 5596 – CNRS / Université Lumière Lyon 2) & Université Rouen (Mehmet-Ali Akıncı, Elif Divitçioğlu, Ali Oker, Iryna Lehka-Lemarchand)

Germany: Universität Potsdam, Zentrum Sprache Variation und Migration (Christoph Schroeder, Meral Dollnick, Christin Schellhardt, Verena Mezger, Ceren Koç-Gültürk, Anne Jähnert, Ginesa Dux, Banu Hueck, Betül Sena Ürkmez, Mathias Schifferings, Florian Kuhn, Patrick Kühmstedt, Esin Işıl Gülbeyaz)

& Freie Universität Berlin, John-F-Kennedy Institute for American Studies (Carol W. Pfaff)

- was supported by the German Research Foundation DFG and the French National Research Agency ANR

Background of the MULTILIT Project

- MULTILIT investigates the language competence of multilingual pupils in Germany (Berlin) and France (Lyon and Paris)

- The project focuses on the family language (mostly Turkish), the societal language German/French and the foreign language English.

- Few studies have investigated the written and oral competence of multilingual immigrant children, related their abilities to their social background and compared them to the abilities of their monolingual fellow students.

- There is a lack of expressive studies concerning the process of third language acquisition by multilingual pupils.

2. Empirical Research Questions

- **How do multilingual pupils develop oral and written competences in their languages?**

    - How, when and under what conditions does academic linguistic competences develop?

    - How does the competences in the different languages correlate with one another?

    - What kind of impact do the different national and social contexts have?

- **Relation between oral and written text products in the different languages**
  - in the different genres
  - with respect to processes of literacy-oriented complexity ("Sprachausbau", cf. Maas 2010)
- **The development of linguistic competences (pseudo-longitudinal)**
  - in the different languages, genres, media
  - with respect to an increase of complexity and diversity: both structurally and in the lexicon
- **Relation between competences in the different languages**
  - with respect to processes of transfer (conceptual, structural) and dynamics of language contact
  - also linked to competence development (see above)

- **The relation between multilinguals' and monolinguals' (German/French/Turkish) text products**
  - with respect to the relation between oral and written text products
  - with respect to the development of linguistic competences
  - with respect to the relation between the school/majority language (German/French) and the foreign language (English)
- **… in correlation with the biographical, sociolinguistic, and sociological data**
  - as collected through the questionnaire
  - as collected through information collected in the schools

3. Methodology

3.1. Project Design

MULTILIT Germany	**MULTILIT France**

• 176 study participants in total
  – 166 German-Turkish bilinguals
  – 10 German monolinguals:
    L1 German (comparison group has to be elaborated on)
 • pupils from different types of school
 (2 primary schools, 2 secondary schools; bilingual and monolingual pupils)
 • Turkish-German bilingual pupils from 4 age groups
 • approximately 20 pupils per age group
	• 163 study participants in total
	  – 84 French-Turkish bilinguals
	  – 79 French monolinguals:
	    L1 French
	
	 • pupils from different types of school :
	 (3 primary schools, 4 secondary schools; bilingual and monolingual pupils)
	 • Turkish-French bilingual pupils from 4 age groups
	 • approximately 20 pupils per age group

⟶ pseudo-longitudinal design

3.2 Data collection

- elicitation technique:
  recalls of a silent video film by Ruth Berman (Tel Aviv, 2002) about problems at school
- After watching the film the pupils produced texts in their familiy language, their language used in school (German or French) and their first foreign language English (only in Germany)
- 2 media: oral, wirtten
- 2 genres: narrative and expository texts
- spoken material: interviews with native speakers
- written material: writing prompts given

MULTILIT Germany: 12 texts per bilingual pupil
                        8 texts per monolingual pupil
                        -> a corpus of **1805 texts**

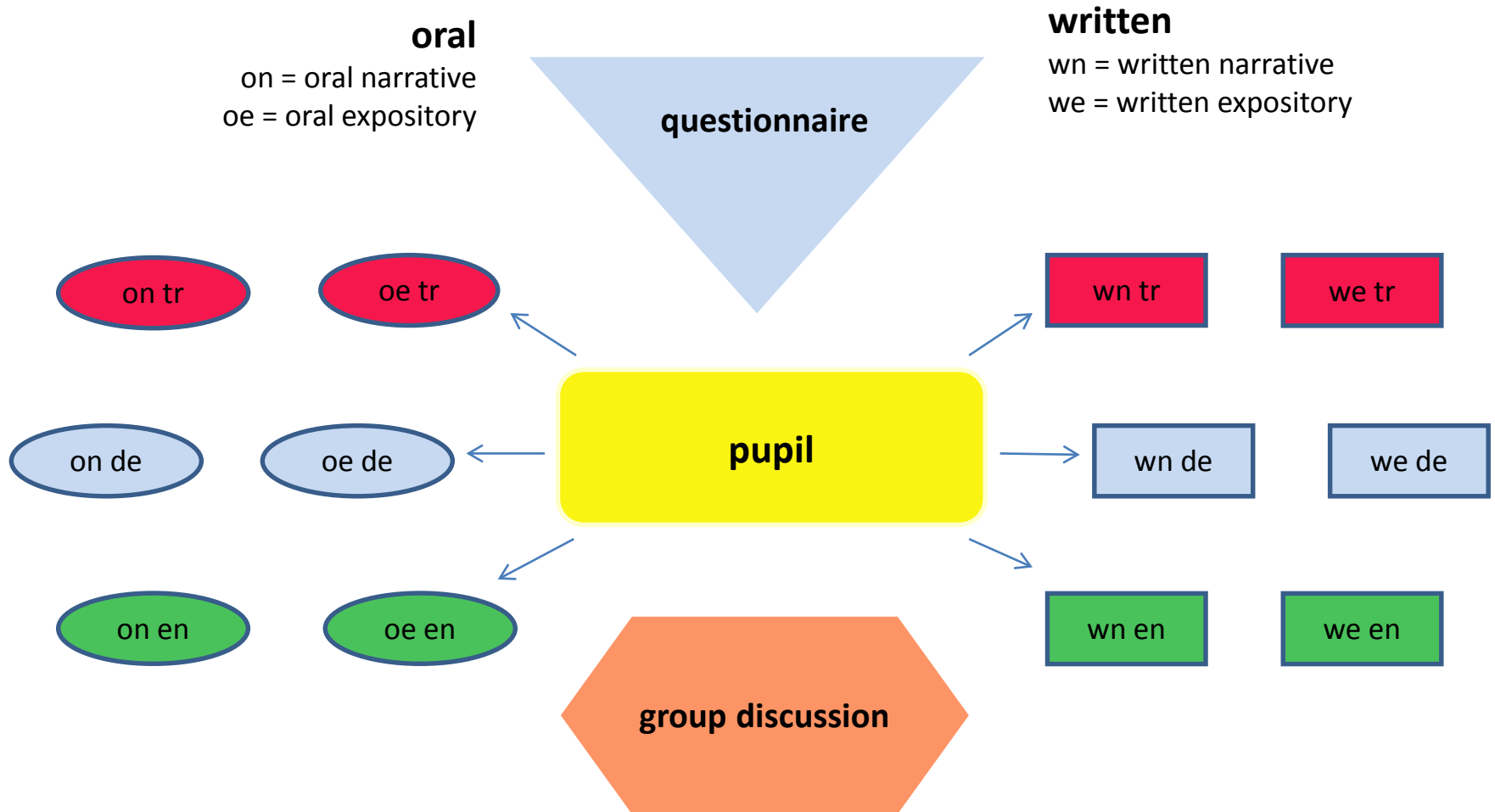**MULTILIT France:**   8 texts per bilingual pupil
                        4 texts per monolingual pupil
                        -> a corpus of **988 texts**

- questionnaire about language background and habits
- Additionally omn Germany: group discussions in all project languages per video
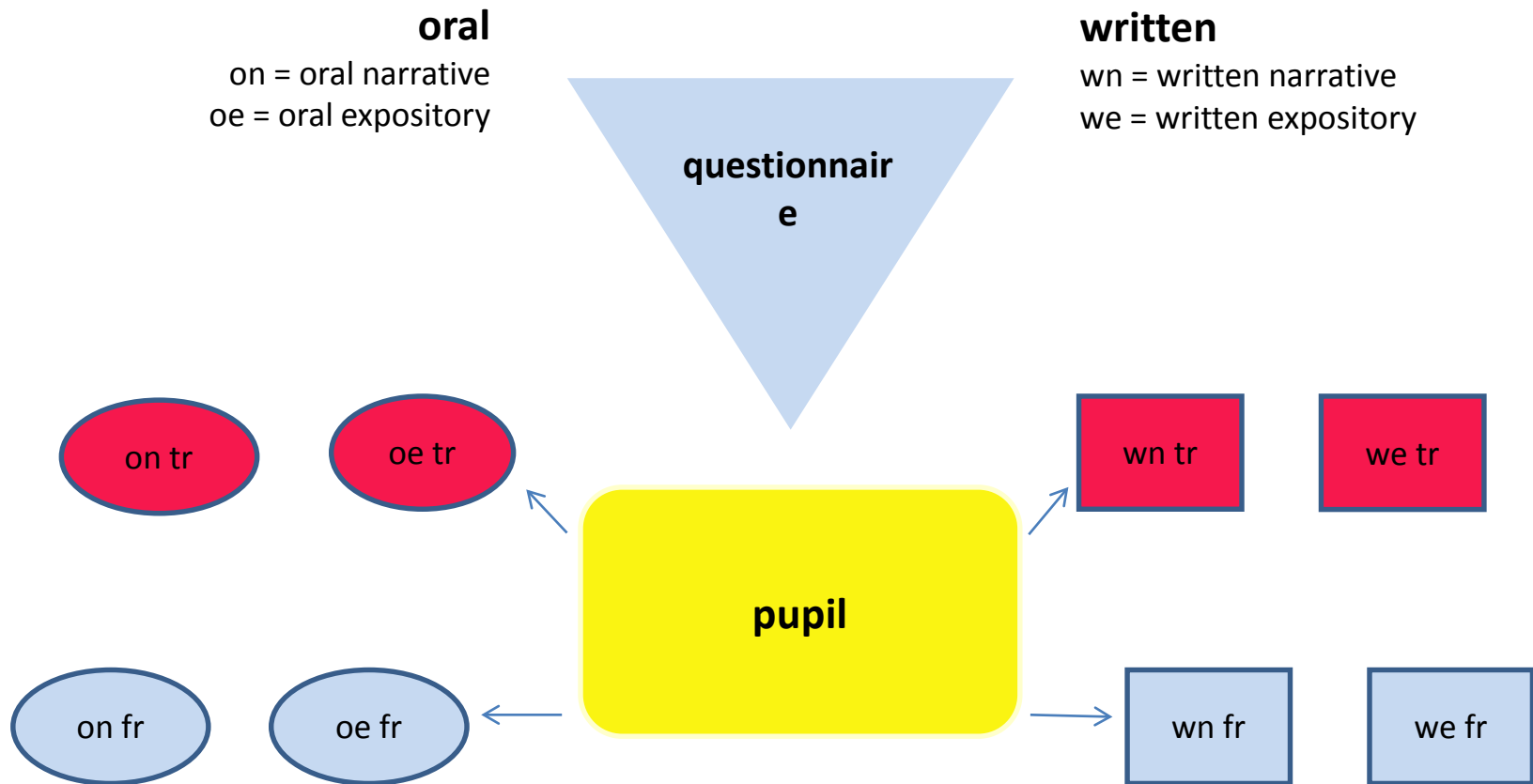
Texts per pupil (MULTILIT Germany)
tr = Turkish, de = German, en = English

oral
on = oral narrative
oe = oral expository

questionnaire

written
wn = written narrative
we = written expository

on tr   oe tr

wn tr   we tr

on de   oe de   **pupil**   wn de   we de

on en   oe en   wn en   we en

group discussion

Texts per pupil (MULTILIT France)
tr = Turkish, fr = French

oral

on = oral narrative
oe = oral expository

written

wn = written narrative
we = written expository

questionnaire

| on tr | oe tr | pupil | wn tr | we tr |

| on fr | oe fr | wn fr | we fr |

3.3 MULTILIT Germany

Processing of collected data : Completed steps

- Transcription of all texts (incl. proof-reading)
  - EXMARaLDA's transcription programme Partitur Editor
  - transcription convention HIAT (incl. changes)
  - segmentation along word boundaries
  - transcription tiers for interviewer and pupil

- development of additional tiers according to analysis criteria in order to carry out quantitive analysis

- development of tiers with target hypotheses for norm-deviating utterances

Processing of collected data: Completed steps

- annotation of subcorpus according to preliminary analysis criteria
  - 28 pupils; 336 texts
  - balanced with regard to age groups, sex, bilingual vs. monolingual instruction
  - only pupils who completed texts in all three languages

- application of POS-Tagging to the complete corpus

- processing of complete corpus in EXMERaLDA's Corpus Manager COMA
  - enables the creation of subsamples

- integration of questionnaire data into the corpus

3.4 Analysis Criteria: A summary

- parts-of-speech (POS)
- noun phrase complexity (NP)
- syntax (SYN)
- complex verbal morphology (MORPH):
   for Turkish and English only
- language mixing (MIX)
- norm deviations (ERR): textual, syntactic, lexical, morphological, orthographical
- textuality: text openings and closings (STANCE and FUNC)
- communicative mode (MODE): monological vs. dialogical
- direct and indirect speech (QT)

3.5 Automatic POS tagging

- We used the automatic part-of-speech tagging software developed by Schmid (1994) for all three languages in the corpus.

- Each element of the verbal tier of the pupils is allocated a part-of-speech tag.

- The POS tagging is based on a set of POS-variables that have been derived from the subcorpus.

3.6 MULTILIT corpus

The MULTILIT corpus contains

- transcriptions of the interview data

- scans of the handwritten texts

- audio recordings of the spoken data

- condensed versions of the data acquired via the questionaire (see manual p. 7-9)

A list of records (publications, MA theses and dissertation projects) from the MULTILIT project can be found at http://www.uni-Potsdam.de/daf/projekte/multilit.html.