Children of Immigrants Longitudinal Survey in Four European Countries

Germany
(CILS4EU-DE)

Technical Report on Accent Measurement

Wave 6 – 2016/17

Version: 7.0.0

Year: 2024

Citation: Weißmann, Markus and Jörg Dollmann. 2024. Children of Immigrants Longitudinal Survey in Four European Countries – Germany. Technical Report on Accent Measurement. Wave 6 – 2016/17, v7.0.0. Mannheim: Mannheim University.

Content

1   Introduction

The accent measurement implemented in the sixth wave of CILS4EU-DE aimed to assess the strength of foreign accents as well as regional dialects among a target population of 19–22-year-old adolescents. We developed the accent measurement in collaboration with phoneticians from the University of Halle-Wittenberg and especially with the help of Prof. Dr. Ursula Hirschfeld.

The accent measurement was carried out in the course of the sixth wave of data collection in the project CILS4EU-DE. The sixth wave consisted of a face-to-face interview as well as an abbreviated mixed-mode follow-up survey for respondents from the initial sample who had not been interviewed in the course of the face-to-face survey. This follow-up survey was not conducted among the refreshment sample, which was additionally sampled in the sixth wave (for more details on the procedure and the differentiation between initial sample and refreshment sample, see Schiel et al. 2016; CILS4EU-DE 2019). Given that the follow-up survey was conducted using telephone, web-based and postal interviews, the accent measurement could not be conducted among this subpopulation. Therefore, it was only conducted among respondents who were surveyed in a face-to-face mode during the sixth wave of data collection and who had consented to being recorded.

The aim of this report is to describe the stimulus material and the recording procedure during the interviews of the sixth wave of data collection of CILS4EU-DE (section 2), the selection and training of the accent raters (section 3), and the subsequent evaluation of the different audio files (section 4). We end this report with a short summary on some descriptive findings, such as the consent rate of the recordings or the amount of missing values on different variables.

2   Stimulus and recording procedure

2.1   The position of the recording in the interview

The recording of the audio sequences took place at the very end of the personal interview, after having collected the panel consent. The accent measurement consisted of two parts: a reading part and a discussion part. In the first part, respondents were asked to read out a short text in order to assess accents during the reading of sentences. To separate reading difficulties from difficulties in pronunciation and accentuation, the interviewer and the respondent held a conversation afterwards. The interviewer recorded the whole procedure. Respondents were told

that the aim of the discussion was to assess the distribution of accents and dialects among the respective age group and that the recording of the talk served to document the feedback. Interviewers were instructed to ask respondents for permission to record the reading of the short text and the subsequent conversation. The consent to the recording was explicitly recorded.[1]

2.2 Reading text

For the first part of the accent recording, interviewers asked the respondents to read out a text from the laptop screen. The reasoning behind this was to make respondents use their language skills in a predefined way and not give them the opportunity to avoid words they did not know or that were difficult to pronounce. In addition, by giving them the task of reading out, respondents had to concentrate their cognitive abilities on a specific task without the possibility of choosing words they were comfortable with. The text 'Goethe und die Studenten' ('Goethe and the students') was chosen due to its structure, which is well suited to reveal speakers' difficulties in pronouncing German sentences in standard German, free of any foreign or regional accents. The text is typically used in the German context to trigger and detect accents. This includes phonological particularities of the German language in terms of segmentals (e.g., correct performance of terminal devoicing, vocals and consonants, or glottal stops) and supra-segmentals (e.g., intonation). It also contains all vocals and consonants with which German language learners usually have problems. Furthermore, the text includes a variety of different syntactic and prosodic structures (e.g., questions, answers, short and long sentences, direct speech etc.), making a good command of the German language necessary to master the reading without an accent.

2.3 Talking and discussion after reading

The interviewers were instructed to engage the respondents in a short conversation about the interview in general after the respondents have finished reading out the text. Respondents were informed that the recording continued so the interviewer would not have to take notes. More specifically, interviewers asked what respondents liked or disliked about the interview, which topics they liked or disliked in the survey, or what recommendations for improvement they would give. This should make respondents engage in extemporaneous speech without consciously adjusting or suppressing their accent.

---

[1] Audio files that did not include such permission or indicated refusal were deleted either beforehand by *infas-Institut für angewandte Sozialwissenschaften* or afterwards by research assistants at the Mannheim Centre for European Research (MZES).

2.4 Execution of the recording

We used built-in microphones of the interview laptops to record the sequences, while respondents read the text from the display with the laptop facing them. Interviewers instructed the respondents in advance, without them seeing see the text on the laptop and thus having no possibility of reading the text silently or aloud beforehand. Consequently, all respondents had the same starting conditions, which should ensure comparability between the readings of different respondents.

3 Accent raters: selection and training

The procedure described above resulted in a total of 4,059 recordings (for the consent rates, cf. Table 1 in section 5). These recordings built the basis for the accent rating, which was carried out by accent raters. In this section, we describe how these accent raters were selected and trained.

3.1 Selection criteria for accent raters

Local students were recruited to administer the evaluations. Vacancies were advertised locally through student mail distributers in different student faculty organisations as well as on student job boards of universities and universities of applied sciences, and of the Federal Employment Agency. Although in practice no particular skills are required to analyse accents and dialects, certain characteristics were preferred, such as a study background in the field of linguistics and philology in general and in German as a second language, German studies, or phonology in particular. In addition, prior experience in phonetic analyses and analyses of voice recordings as well as knowledge of the International Phonetic Alphabet (IPA) and fundamental phonetic terms were desired. German as a first language or native-like proficiency was a necessary prerequisite.

Ultimately, two rounds of three accent raters each were employed,[2] who studied in the fields of English studies, German studies, North American studies, Hispanic Studies, Japanese Studies, German as a second language, translation studies, language studies, and cultural studies.[3] The

---

[2] While the first three assistants were employed in August 2016, a second round of three students were employed in October 2016 to support the staff (one rater quit before the end of the evaluation procedure). The training procedure was identical for both groups. The total duration, including training and accent coding, was from August 2016 to July 2017.

[3] Translations-, Sprach- und Kulturwissenschaften; Anglistik; Germanistik; Deutsch als Zweitsprache; Japanologie; Amerikanistik; Hispanistik.

raters were all female, stemmed from Germany, Kazakhstan and Paraguay, and were between 21 and 29 years old.

3.2   Training of accent raters for the analyses

All accent raters underwent a two-day long introduction to the field of phonetics and phonology at the University of Mannheim, carried out by a trained lecturer for linguistics and phonetics from the University of Halle-Wittenberg. Course contents included an introduction to the different segmentals and supra-segmentals of speech as well as to causes and consequences of foreign accents. In addition, all accent raters underwent transcription exercises using the IPA. This theoretical introduction gave raters a basic understanding of the segments of spoken language, how spoken language is built, how sounds are produced, what difficulties speakers from different linguistic backgrounds encounter specifically when speaking standard German, and why this is the case. Thus, the introduction was designed to help raters detect problems, inadequacies, and deviances from standard German in speech. Besides this theoretical background introduction, raters carried out evaluation exercises. In a first step, sample recordings with speakers from different countries and with different strengths of accent were presented and discussed in the group.[4] In a further step, samples from the CILS4EU-DE survey were analysed jointly, using the final evaluation sheet.

3.3   Preparatory phase, comparability, and practice

To ensure comparability between the raters, we decided to adopt measures that would help them train their ability to detect accents, grade their strengths, and share their experiences. Thus, before analysing the voice recordings, accent raters were given at least one month of exercises and training. During this time, raters were asked to analyse recordings jointly and discuss their individual evaluation procedure. During the first meetings, raters listened to recordings together, discussed issues during the recordings, and worked out ratings for the individual respondents. The goal was to allow raters to develop a common scheme of evaluation by jointly elaborating ratings. While the first sessions exclusively consisted of these meetings, the frequency was reduced over time, and raters began to evaluate recordings individually and gathered only weekly to biweekly. In those later meetings, raters collected recordings that were special in a certain way and discussed these issues. This included above-average ratings, special types of accents and dialects, or recordings in which evaluations were not straightforward.

---

[4]   Sample recordings were taken from Deutsche Welle – Deutsch Lernen – Lernerporträts (http://www.dw.com/de/lernerportr%C3%A4ts-archiv/a-5465885 ; accessed 30.05.2018).

These joint sessions continued also during the phase of analysis, although then on an occasional basis. For an overview of the number of recordings evaluated by each rater, see Table 1 in section 5.

4 Evaluation: Information gathered from the recordings

In this section, we describe the information that was gathered from the recordings. The rating sheet after which the raters evaluated the recordings can be found in the Appendix.

4.1 The main aspects: foreign and regional accents

The main aim of the evaluation was to assess the degree to which foreign and regional accents are distributed among young adults with and without migration background in Germany. Consequently, rating scales that allowed the raters to evaluate a respondent's accent strength made up the core part of the evaluation sheet. Both foreign and regional accents were coded on a 9-point scale, reaching from 1 ('no accent') to 9 ('very strong accent'), with only the extreme categories being labelled. Respondents for which raters did not detect any accent at all were coded with 1 ('no accent') while the remaining eight points were used to rank the degree of accent strength. Consequently, the lowest possible accent strength was represented by the value 2, which was used as the minimum for any detectable accent in a respondent's speech. Following the structure of the interview (see section 2), raters had to evaluate accents first in reading and then in talking.

Additionally, for both modes of communication, raters evaluated the comprehensibility of the speaking material in two ways. First, they rated the general comprehensibility on a 6-point scale from 'very good' to 'very bad'. Second, reasons for impaired comprehensibility, such as accent strength, stuttering, or mumbling, were documented. For both ratings, raters were instructed to take into account only communicational and language aspects but not technical or environmental distractions, such as background noises or the quality of the recording.

For all questions, raters could also tick a box to indicate that an evaluation was not possible for technical or practical reasons. This applied to situations in which recordings were terminated before the respective part could be administered (e.g., rating of accents in extemporaneous speech when recordings were terminated during or directly after the reading part) or in situations in which technical issues made an evaluation impossible (e.g., loud background noises, microphone problems, etc.).

4.2 Additional information on situation and conversation

To get a better idea of the language abilities, the background of the respondent, and the situation in which the interview took place, additional information about the respondent and the interview situation was collected from the raters.

4.2.1 Background of the respondent

Several aspects of the respondents' language abilities and background were collected. First, raters assessed the first language of a respondent. Raters could report up to five different languages. If raters could not identify the first language, 'unclear' was recorded.

Second, based on the regional accent, raters then established the German region from which a respondent came. If raters were unsure about the respondent's exact linguistic origin, they chose the category 'region unclear'. If the respondent's language did not contain any indicators of a specific regional origin, raters chose the category 'no region'. In addition, raters could make further comments about the presumed region.

Third, raters assessed the respondents' use of verbal means during the conversation on a 6-point scale, ranging from 'very good' to 'very bad'. Four aspects of conversational skills were rated: grammar, vocabulary, expression, and appropriateness. If verbal means could not be assessed, for instance due to low talkativeness, raters indicated 'cannot be assessed'.

Fourth, raters should determine the respondent's level of general education, choosing between 'degree from lower secondary school (Hauptschulabschluss)', 'degree from intermediate secondary school (Realschulabschluss)' and 'degree from upper secondary school (Abitur)'. If the level of education could not be assessed, raters indicated 'cannot be assessed'.

Again, for all questions, raters could also tick a box to indicate that an evaluation was not possible for technical or practical reasons.

4.2.2 Interview situation

To better capture the interpersonal situation in which the conversation took place, several aspects related to the persons involved were assessed. First, raters graded how talkative the respondent was, choosing between 'very talkative', 'rather talkative', 'not very talkative', and 'not at all talkative'. This variable was used to assess how much talking material was available for raters to deduce information on the accent strength and use of verbal means. In addition, the personal relationship between interviewer and respondent was assessed, selecting 'very good', 'good', 'average', or 'less well'. Lastly, also the strength of foreign and regional accent of the interviewer was evaluated analogously to the ratings of the respondents.

Again, for all questions, raters could also tick a box to indicate that an evaluation was not possible due to technical or practical reasons.

4.2.3 Other information

Lastly, information on the recording quality (on a 6-point scale ranging from 'very good' to 'very bad') as well as information on whether or not the recording was terminated prematurely was included. If the recording was terminated prematurely, the time of the termination was indicated (before reading, during reading, between reading and conversation, during conversation). This information can be used to better grasp the length of the recording and the information from which raters could generate their assessments.

4.3 Data entry mode

All evaluations were entered with the help of the online survey platform unipark by QuestBack. Admission to the evaluation survey depended on entering a unique code, which was set to be the respondents' ID and ensured the exact matching of recording, evaluation, and respondent ID.

5 Descriptive overview

This descriptive overview gives readers an impression of the data in terms of participation, missing information, premature termination, and distribution of raters over recordings. The denominator for the total participation rate in the accent recordings is the number of interviews in the face-to-face mode (long version). The denominator for all other information is the total number of recordings.

Table 1: Information on accent data

Consent to recording	4,059 of 5,074 (79.99 per cent)
Missing evaluation due to technical or practical reasons or premature termination	
Accent rating	
Reading	18 (0.44 per cent)
Talking	237 (5.83 per cent)
Reading or talking	242 (5.96 per cent)
Assessment of verbal means	185 (4.55 per cent)
Premature termination of recording	
Before reading	5 (00.12 per cent)
During reading	1 (00.02 per cent)
Between reading and conversation	215 (05.29 per cent)
During conversation	35 (00.86 per cent)
Overall	256 (06.30 per cent)
Number of records per rater	
Rater 1	1028 (25.32 per cent)
Rater 2	1106 (27.24 per cent)
Rater 3	1109 (27.32 per cent)
Rater 4	526 (12.95 per cent)
Rater 5	290 (07.14 per cent)

6 References

Schiel, Stefan, Katharina Sandbrink, Folkert Aust, Nina Chudziak, and Martin Kleudgen, 2016. **Methodenbericht. CILS4EU-Projekt: CAPI-Befragung von jungen Erwachsenen in Deutschland**. Bonn: infas Institut für angewandte Sozialwissenschaft GmbH.

CILS4EU-DE. 2019. C**hildren of Immigrants Longitudinal Survey in Four European Countries – Germany. Technical Report. Wave 6 (short version questionnaire) – 2016/17**, v4.0.0. Mannheim: Mannheim University.

7 Appendix

Evaluation sheet used by raters

Accent rating: reading

D1: **Rating of *foreign* accent: <u>read text</u>**
Scale: Strength of accent

no accent								Very strong accent	Not possible (tech. issues)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

D2: **Rating of *regional* accent: <u>read text</u>**
Scale: Strength of accent

no accent								Very strong accent	Not possible (tech. issues)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

D3: Comprehensability of <u>read text</u>
Not only regarding strength of foreign and regional accent, but also regarding stuttering, mumbling, talking too fast, and the like – but **not** regarding quality of recording!

Very good					Very bad	Not possible (tech. issues)
☐	☐	☐	☐	☐	☐	☐

D4: Comprehensability of <u>read text</u> is impaired by:
Please do not note background noises or quality of recording! Multiple answers possible

- Strength of foreign, regional accent or sociolect
- Mumbling
- Stuttering
- Lisping
- Cluttering
- Stammering
- Lalling
- Filler sounds/words („er", „um", etc.)
- Loudness: too loud
- Loudness: too quiet
- Speed: too fast
- Speed: too slow
- Other reason: _____
- No impairment
- Not possible (technical issues)

Accent rating: conversation

D5: Rating of *foreign* accent: talking and conversation
Scale: Strength of accent

no accent								Very strong accent	Not possible (tech. issues)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

D6: Rating of *regional* accent: talking and conversation
Scale: Strength of accent

no accent								Very strong accent	Not possible (tech. issues)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

D7: Comprehensability of talking and conversation
Not only regarding strength of foreign and regional accent, but also regarding stuttering, mumbling, talking too fast, and the like – but **not** regarding quality of recording!

Very good					Very bad	Not possible (tech. issues)
☐	☐	☐	☐	☐	☐	☐

D8: Comprehensability of talking and conversation is impaired by:
Please do not note background noises or quality of recording! Multiple answers possible

- Strength of foreign, regional accent or sociolect
- Mumbling
- Stuttering
- Lisping
- Cluttering
- Stammering
- Lalling
- Filler sounds/words („er", „um", etc.)
- Loudness: too loud
- Loudness: too quiet
- Speed: too fast
- Speed: too slow
- Other reason: _____
- No impairment
- Not possible (technical issues)

Information on the respondent

D9: Presumed first language
(also multiple answers possible)

- First language 1: _____
- First language 2: _____
- First language 3: _____
- First language 4: _____
- First language 5: _____
- First language unclear
- Not possible (technical issues)

D10: Presumed region in Germany
(Locality, region, or language region)

- Low German
- West Central German
- East Central German
- West Upper German
- East Upper German
- Region unclear
- No Region
- Not possible (technical issues)

D11: More detailed information on the presumed region: _____

D12: Assessment of verbal means during <u>conversation</u>
Consider anything noteworthy (grammar, vocabulary, expression, situational-stylistic appropriateness)

	very good					very bad	cannot be assessed	not possible (tech. issues)
Grammar	☐	☐	☐	☐	☐	☐	☐	☐
Vocabulary	☐	☐	☐	☐	☐	☐	☐	☐
Expression	☐	☐	☐	☐	☐	☐	☐	☐
Appropriateness	☐	☐	☐	☐	☐	☐	☐	☐

D13: Presumed level of education

- Degree from lower secondary school (Hauptschulabschluss)
- Degree from intermediate secondary school (Realschulabschluss)
- Degree from upper secondary school (Abitur)
- Cannot be assessed
- Not possible (technical issues)

Interview situation

D14: **How talkative was the respondent?**

very talkative	rather talkative	not very talkative	not at all talkative	Not possible (tech. issues)
☐	☐	☐	☐	☐

D15: **Relationship between the respondent and the interviewer**

very good	good	average	less well	not at all good	Not possible (tech. issues)
☐	☐	☐	☐	☐	☐

D16: **Rating of *foreign* accent: <u>interviewer</u>**
Scale: Strength of accent

no accent								Very strong accent	Not possible (tech. issues)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

D17: **Rating of *regional* accent: <u>interviewer</u>**
Scale: Strength of accent

no accent								Very strong accent	Not possible (tech. issues)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

D18: **Quality of recording (technical):**

very good					very bad
☐	☐	☐	☐	☐	☐

D19: **Was the record terminated prematurely?**

- No termination
- Yes, abortion before reading
- Yes, abortion during reading
- Yes, abortion between reading and conversation
- Yes, abortion during conversation