

## Assessment structure

Assessment in sign language teaching is divided into the three competence areas of reception, production and interaction and these are based on the descriptors of the levels of the CEFR.

### Possible structure of testing procedures in sign languages

The use of e-learning as a tool is ideal for the design of test procedures – offering many possibilities with regard to tasks and test forms. Thus, e-learning opens up a range of possibilities for designing test tasks in different ways. It is advisable to have used e-learning already in class before testing in order to give learners the opportunity to practise different tasks. What is more, different competence areas can be examined using different tools and task forms. Learners should be given an opportunity to familiarise themselves with the e-learning platform in order not to be confronted with additional technical challenges or uncertainties during assessment. This significantly reduces stress during the exam.

By using an e-learning approach it is possible to create a transparent and, above all, objective assessment of performance to give a binary result: right or wrong. This minimises any subjective influence on the assessment. Moreover, the same conditions (time, task, assessment) are applied to all learners during the test. Another advantage is the automatic display of errors and improvements for learners.

There are various methods available for the task forms in the different areas of competence (reception, production, interaction). Depending on the level of competence and requirements, suitable tasks can be selected and the test procedure can be designed accordingly (cf. Barbeito Rey-Geißler / Geißler 2018).

Examples of test methods for the reception competence:

- “drag and drop” marking
- specific assignment of statements to a picture
- writing a short individual answer
- clicking to decide “right or wrong”
- multiple choice tasks



Examples of test methods for production: video recording

- presenting before the plenum

Examples of test methods for interaction:

- group discussion
- partner discussion



**Possible reception task:**

**(competence level A1 according to the CEFR)**

| Competence level | Discretionary descriptors  | Structure of the testing procedure (indications)  | Examples   |
|------------------|--|---|--|
| A1               | <ul style="list-style-type: none"> <li>- Can understand expressions relating to his/her own environment and simple short sentences</li> <li>- Can follow important and familiar clues</li> </ul> | <ul style="list-style-type: none"> <li>- Short and simple sign-language statements</li> <li>- Visual design of the test procedure using symbols or images to train visual perception</li> <li>- Reduction of written explanations and increased visual presentation of tasks and answering options</li> </ul> | <ul style="list-style-type: none"> <li>- Description of clothes (forms, colours, patterns, size)</li> <li>- Description of people (features such as eyebrow, hair, facial shapes)</li> <li>- Times</li> <li>- Calendar / weekly schedule (time reference and announcing a deadline or an appointment)</li> </ul> |

**Description of clothes:**

In this task, a short sign-language video might show a specific feature (form, colour, pattern, size). Afterwards, learners should – for instance, in a multiple-choice task – click the right answer from several options. Drawings of the various features can be displayed as answering options (thus no written description) (see fig. 1).



(Figure 1)



Description of people:

Here, the “drag and drop” method might be a good idea. In a sign-language video, the appearance of a person could be described and their name and age could be mentioned. Learners are shown drawings of three different people as possible answers, to whom they should assign the correct name and age (fig. 2).



(Figure 2)

Times/calendar:

The learners are shown a calendar overview with various appointments (“My weekly schedule”, fig.3).

The video shows a sign language statement about a specific event (e.g. “Tomorrow morning I’m going to school/university”). The learner then has to assess in terms of “true/false” whether the statement corresponds to the displayed calendar sheet (if tomorrow’s calendar sheet indicates for the morning that shopping is the order of the day, then the statement will have to be assessed as false). This approach checks whether the learner has properly understood the content.

|            | Montag<br>19. Februar  | Dienstag<br>20. Februar | Mittwoch<br>21. Februar      | Donnerstag<br>22. Februar | Freitag<br>23. Februar | Samstag<br>24. Februar | Sonntag<br>25. Februar |
|------------|------------------------|-------------------------|------------------------------|---------------------------|------------------------|------------------------|------------------------|
| Früh       | [Icon: person walking] | [Icon: person walking]  | [Icon: person in wheelchair] | [Icon: person walking]    | [Icon: person walking] | [Icon: person walking] | [Icon: person walking] |
| Vormittag  | [Icon: person at desk] | [Icon: person at desk]  | [Icon: person at desk]       | [Icon: person at desk]    | [Icon: person at desk] | [Icon: person at desk] | [Icon: person at desk] |
| Mittag     | [Icon: person at desk] | [Icon: person at desk]  | [Icon: person at desk]       | [Icon: person at desk]    | [Icon: person at desk] | [Icon: person at desk] | [Icon: person at desk] |
| Nachmittag | [Icon: person at desk] | [Icon: person at desk]  | [Icon: person at desk]       | [Icon: person at desk]    | [Icon: person at desk] | [Icon: person at desk] | [Icon: person at desk] |
| Abend      | [Icon: pizza]          | [Icon: person at desk]  | [Icon: person at desk]       | [Icon: person at desk]    | [Icon: person at desk] | [Icon: person at desk] | [Icon: person at desk] |
| Nacht      |                        | [Icon: person at desk]  | [Icon: person at desk]       | [Icon: person at desk]    | [Icon: person at desk] | [Icon: person at desk] | [Icon: person at desk] |

Select one:  
 True  
 False

## Possible reception task

(competence level B2 according to the CEFR)



| Competence level | Discretionary descriptors   | Structure of the testing procedure (indications)   | Examples  |
|------------------|---|--|---|
| B2               | <ul style="list-style-type: none"> <li>- Can understand statements relating to his or her own environment as well as everyday and socially relevant topics</li> <li>- Can understand discussions and conversations</li> </ul> | <ul style="list-style-type: none"> <li>- Having sign language statements on realistic everyday topics performed by native users</li> <li>- Possible reduction of memorising as part of the task through multiple choice</li> </ul> | <ul style="list-style-type: none"> <li>- Narrative about the recent culture days</li> </ul> |

### Possible examples:

Learners see a sign language narrative, followed by a variety of content questions in multiple-choice format that they are supposed to answer.

At first, learners see the whole video to get a first overview of the context. Afterwards the video is divided into individual clips; learners watch the clips and are then asked to answer a question by selecting one of the multiple-choice answers (fig. 4).

The answering options are also short sign-language clips featuring only one correct answer each. The sign-language expression used is not the same as in the test video; the content of that video is merely summarised. This helps to make sure that only the correct reception in terms of content is assessed, regardless of the sign-language style and execution.

(Figure 4)



## Possible production task:

(competence level A1 according to the CEFR)



| Competence level | Discretionary descriptors  | Structure of the testing procedure (indications)  | Examples   |
|------------------|--|---|--|
| A1               | <ul style="list-style-type: none"><li>- Can produce simple mainly isolated phrases about people and places.</li><li>- Can describe him/herself, what he/she does and where he/she lives.</li></ul> | <ul style="list-style-type: none"><li>- Task provided in the respective native language</li><li>- Video recording of 3 to 5 minutes maximum sufficient</li><li>- Using visual media and recording the utterance as a video</li><li>- Preparing a monologue in the target language</li></ul> | <ul style="list-style-type: none"><li>- Tell us about your environment.</li><li>- Present your family relations.</li></ul> |

### Video recording:

A video recording is made after a specific topic has been provided.

### Presenting before the plenum:

Alternatively, the utterance may be presented to a plenum. Care should be taken, however, that the learners are not overwhelmed by the additional requirement of a presentation in front of a large number of people some of whom they might know. This should be seen in relation to the CEFR and adapted accordingly.

Following the CEFR, the language products are evaluated according to defined criteria: range, correctness, fluency, coherence. The range covers everything from the use of learning materials to the use of language in a group. The correctness criterion refers to the sentence structure, and it is evaluated to what extent sentence structures are internalized and the grammar is executed correctly. For fluency the execution of sign language is evaluated. Beginners still sign rather timidly and falteringly, advanced learners manage to express themselves more self-confidently and fluently. Coherence refers to the linking of content and the associated logical expressions.



When assessing the production at A1 level, the focus is on vocabulary and grammatical structures. During the evaluation, tables can be used to check which elements were used during production.

As is done with testing the CEFR, utterances are assessed using the following criteria. The respective weighting is given as a percentage in brackets:

For A1:

- range (30%)
- correctness (30%)
- fluency (20%)
- coherence (20%)



The weighting shows where the focus lies at the respective competence level.

The results are converted into percentages and an individual grade is calculated from the total percentage. It is possible to look at individual performance and compare it to the average of the group (cf. appendix 1).



**Possible production task:**

**(competence level B2 according to the CEFR)**

| Competence level | Discretionary descriptors  | Structure of the testing procedure (indications)  | Examples   |
|------------------|--|---|--|
| B2               | <ul style="list-style-type: none"> <li>- Can hold a subject-specific conversation</li> <li>- Can understand complex topics in detail</li> <li>- Can express him- or herself clearly and in detail in technical terms</li> <li>- Can provide information on current events</li> </ul> | <ul style="list-style-type: none"> <li>- Having sign language statements on realistic everyday topics performed by native users</li> <li>- Reduction of memorising as part of the task</li> <li>- Providing two short key words or buzz words relating to a specific topic</li> <li>- Video recording of 5 to 8 minutes maximum sufficient</li> </ul> | <ul style="list-style-type: none"> <li>- Recent elections in Germany</li> <li>- Security issues in Germany</li> <li>- Culture days of the deaf</li> <li>- Deaf festival</li> </ul> |

### Possible procedure:

Providing two topics. After the learner has chosen a topic, the task is explained in more detail. Here, learners should take the subject-specific aspect into account. Examples:

1. Recent elections in Germany
2. Security issues in Germany



- 1) Please describe the chronological procedure for the recent election, explaining your view of the election and how it went. Use arguments to support your opinion.
- 2) Outline your position on security in Germany and make suggestions for possible changes.

The language outputs are evaluated according to defined criteria as is done with CEFR assessment. In addition to the criteria that are already used in the evaluation of A1 production, at language level B2 specific aspects are evaluated; it is essential to be able to express oneself in sign language on specialist topics and to explain different issues. In addition, there is assessment of the ability to express arguments and explain personal opinions. There is therefore a specific assessment, not referring to everyday conversations, as is the case of the other four criteria.



In line with the CEFR, the utterances are assessed using the following criteria. The respective weighting is given as a percentage in parentheses:

For B2:

- subject-specific aspects (25%)
- range (15%)
- correctness (15%)
- fluency (20%)
- coherence (25%)



The weighting shows where the focus lies at the respective competence level.

The results are converted into percentages and an individual grade is calculated from the total percentage. It is possible to look at individual performance and compare it to the average of the group (cf. appendix 2).

## Assessment of sign-language interaction

The situation should be as authentic, direct, and realistic as possible. Here, realistic situations can be created with the CEFR competence levels in mind.



The conversation should be hosted by a third person (examiner), who not actively influence it. He or she should control the conversation to make sure all the participants get opportunities to express themselves. This can be done, for example, by means of signals such as directly addressing the person that has said only little so far by asking them for their opinion, or by means of controlling the conversation if the interacting person deviates from the topic. Participants could also be directly asked to contribute to the conversation and express their point of view and, in this way, actively participate in the discussion. An appropriate time frame should also be set. It is a good idea to record the interaction with cameras to make sure that different perspectives are taken into account and also that the conversation can be accessed later on for assessment or review purposes.

### **Possible interaction task:**

**(competence level A1 according to the CEFR)**



| Competence level | Discretionary descriptors   | Structure of the testing procedure (indications)   | Examples   |
|------------------|---|--|--|
| A1               | <ul style="list-style-type: none"> <li>- Can ask how people are and react to news.</li> <li>- Can understand everyday expressions aimed at the satisfaction of simple needs of a concrete type, delivered directly to him/her in clear, slow and repeated sign by a sympathetic signer.</li> <li>- Can indicate time by use of such lexicalised phrases like NEXT WEEK, RECENT NOVEMBER, 3 O'CLOCK, etc.</li> </ul> | <ul style="list-style-type: none"> <li>- max. 2 persons interacting</li> <li>- Video recording of 5 to 10 minutes maximum</li> </ul> | <ul style="list-style-type: none"> <li>- Tell each other about your family relations</li> <li>- Discuss your individual weekly schedules.</li> <li>- Tell each other about the organisation for a next party.</li> </ul> |

When assessing interaction, the focus is not on the smoothness or speed of the conversation. The focus is on the extent to which questions and comments can be formulated, understood and appropriately answered in order to achieve a coherent conversation. The assessment of interaction at A1 level therefore focuses on understanding and conversational skills.

In line with the CEFR, the interactions are assessed using the following criteria. The respective weighting is given as a percentage in parentheses:

For A1:

- interaction (20%)
- range (20%)
- correctness (20%)
- fluency (20%)
- coherence (20%)



The weighting shows where the focus lies at the respective competence level.

The results are converted into percentages and an individual grade is calculated from the total percentage. It is possible to look at individual performance and compare it to the average of the group (cf. appendix 3).

### Possible interaction task:

#### (competence level B2 according to the CEFR)

| Competence level  | Discretionary descriptors   | Structure of the testing procedure (indications)   | Examples  |
|---|---|--|---|
| B2<br><br> | <ul style="list-style-type: none"> <li>- Can express his/her ideas and opinions with precision, and present and respond to complex lines of argument convincingly.</li> <li>- Can use the language fluently, accurately and effectively on a wide range of general, academic, vocational or leisure topics, marking clearly the relationships between ideas.</li> </ul> | <ul style="list-style-type: none"> <li>- max. 4 to 6 interactants</li> <li>- Time frame of the video recording depending on group size (4 persons = 20 minutes)</li> <li>- Discussion using pros and cons</li> </ul> | <ul style="list-style-type: none"> <li>- Specialist subject</li> <li>- Everyday politics</li> <li>- Use of social media (e.g. Facebook) and what everybody thinks about it</li> <li>- A current and much discussed topic in the deaf community</li> </ul> |

When assessing interaction at B2 level, a group of participants between 4-6 persons is evaluated. In the process, there is assessment of how the individual participants manage to take an active part in the conversation and which strategies they use in order to contribute a comment or take over the conversation at the appropriate time without simply interrupting the conversation. In addition, there is evaluation of extent to which participants are able to link in with the current topics of the discussion and to deepen these further, also using various sources.

A second evaluation criterion for interaction at B2 level is turntaking. The strategies participants use to keep or give up their turn, to produce a fluent conversation are observed. A further criterion is the extent to which participants concentrate on the conversation in spite of surrounding, distracting visual stimuli. The concentration of the participants on the conversation despite surrounding, distracting visual stimuli is also evaluated as is the extent to which they can adapt their register to the respective degree of formality (e.g. not giggle in a serious discussion).

The assessment focuses on these factors during interaction at B2 level, with grammatical aspects included as a secondary aspect.



In line with the CEFR, participation in the conversation is assessed using the following criteria. The respective weighting is given as a percentage in parentheses:

- interaction (25%)
- range (15%)
- correctness (15%)
- fluency (20%)
- coherence (25%)



The weighting shows where the focus lies at the respective competence level.

The results are converted into percentages and an individual grade is calculated from the total percentage. It is possible to look at individual performance and compare it to the average of the group (cf. appendix 4).