OULUN AMMATTIKORKEAKOULU

Mikhail Demidov

# TEACHING INTERPRETATION: PROGRESS ASSESSMENT TOOL

A portfolio-based study of the prototypes of the peer-assessment tool for interpreter training

# TEACHING INTERPRETATION: PROGRESS ASSESSMENT TOOL

A portfolio-based study of the prototypes of the peer-assessment tool for interpreter training

Mikhail Demidov
Thesis
Spring 2022
Master in Education Entrepreneurship
Oulu University of Applied Sciences

## ABSTRACT

Oulu University of Applied Sciences
Master in Education Entrepreneurship, Thesis

Author: Mikhail Demidov
Title of Master's Thesis: Teaching interpretation: progress assessment tool
Supervisor: Pablo Santur
Term and year of completion: Spring 2022 Pages: 125

The experience of multilingualism at the United Nations (UN) and practices of the International Association of Conference Interpreters (AIIC) have shaped useful hands-on approaches to teaching translation and interpreting (T/I). The Interpretive Theory of Translation proposed by AIIC's Danica Seleskovitch has become a valuable pedagogical tool that remains the mainstay of top-notch European programmes in T/I. Meanwhile, performance assessment in the routine T/I teaching is far from being uniform. Many professional communities and institutions are using their own sets of criteria, with evaluation sheets being the main type of tool. A CEFR-like competence framework was proposed by PACTE group from Autonomous University of Barcelona. However, it applies solely to translation, but not to interpreting.

This study seeks to identify the characteristics of a software tool for developing the invariant competencies that are common for consecutive and simultaneous interpretation. The thesis focuses on the competencies that such tool should support, and the characteristics of the user interface. Qualitative approach is implemented: the study relies on the results of a focus group of professional interpreters, a workshop for educators at Oulu University of Applied Sciences and an interview with a freelance interpreter with experience at the UN specialized agencies. The study is portfolio based: the author designed three different Prototypes of a tool for interpreter training and presented them during the meetings. The participants were asked to assess the interpreter performance using one or several Prototypes and talk about their experiences. Some Prototypes had pre-defined lists of criteria and offered a particular assessment framework. Other Prototypes did not limit the participants in any way. The meetings were recorded and transcribed. Content analysis of the recordings was carried out.

The main results showed that professional interpreters preferred not to have any scaffolding and needed an adjustable tool (e.g. allowing to assign relative weights to the criteria). The invariant criteria included the accuracy of messaging, mastery of the target language, and coherence of presentation. This study may be useful for the development of tools for interpreter training, as well as competence frameworks in interpretation.

Keywords: Interpreter training, Structured assessment, Competence framework, CEFR, PACTE, Epistemic profiles, Prototype.

## CONTENTS

# 1. INTRODUCTION

The author of this study is a professional conference interpreter/translator and CELTA-certified English language instructor with over 10 years' entrepreneurial experience. He also has academic background teaching English for Specific Purposes (ESP) to graduates and Consecutive Interpreting to postgraduates at the university level. His interests include the construction and activation of background knowledge (Demidov, 2019), and effective delivery in speech (Demidov, 2020). He advocates the active usage of schemes, algorithms and criteria, as well as the structured assessment of learner performance by peers, tutors and learners themselves. The author has proposed a progress assessment toolkit (Demidov, Teaching Interpretation: Progress Assessment Toolkit (English version), 2020) which he has used in various forms since 2016. In this portfolio-based thesis the author would like to present and examine three Prototypes of the peer-assessment tool. In the long run, the tool is meant to become a part of an expert system which will be used:

a) as a virtual environment (medium) for teaching interpretation to degree students;
b) as a system to arrange demand-based training for professionals (e.g. in continuing professional development);
c) as a means to identify learning difficulties and provide practical advice;
d) as a means to work on individual skills, some of which are rarely addressed in isolation in the Translation curricula (e.g. psychophysiological skills).

The thesis explores the characteristics of a software tool for developing interpreter competence in consecutive interpretation (or, to be more precise, the core competencies that are common for consecutive and simultaneous interpretation, see section 2.3. Development of Translation Competences in the European Region). The thesis will focus on the competencies that such tool should support, and the characteristics of the user interface[1]. When referring to a particular Prototype, the author uses numerical indexes, i.e. 'Prototype 1', 'Prototype 2', and 'Prototype 3'. With the exception of Chapter 3 that describes the existing software from other manufacturers, the author uses the term 'tool' generically, thus referring to a hypothetical final product of his prototyping. This final tool may be an updated/the existing version of Prototypes 1 – 3, their combination, or another iteration based on them.

[1] User interface (UI) is defined as "the physical and digital components that allow user to communicate with a machine or device <…> UI refers to both the physical and digital (on-screen) interactions between a human user or operator and a device or piece of equipment" (Swan, 2015, Chapter 23, p.10).

# 2. THEORETICAL BASIS

This section describes the principles of professional practice and their links with the similar trends elsewhere. It also addresses the education theories that lie at the heart of interpreter training in Europe and analyzes an example of a curriculum in consecutive interpretation.

## 2.1. Principles of the professional practice

Routledge Encyclopedia of Translation Studies defines interpreting as "the oral or signed translation of oral or signed discourse, as opposed to the oral translation of written texts" (Baker, 2008, p. 51). Conference interpreting is one of the forms of interpreting, and one of its distinctive features is "high performance level, the latter as described in particular by AIIC"[2] (Baker, 2008, p. 52). The experience of multilingualism at the United Nations (United Nations, 2020) and the Institutions of the European Union have shaped useful practical approaches and promoted the development of a number of widely recognized training programmes in conference interpreting (United Nations, 2020; AIIC, 2004). For example, European Masters' Programme in Conference interpreting (EMCI) began in 1997 to cover the demand for greater numbers of interpreting staff (Baker, 2008). Such programmes are compliant with AIIC's Best Practice Criteria (AIIC, 2021).

This means, inter alia, that:

- the courses are available solely for post-graduate students;
- the curriculum includes training in both simultaneous and consecutive interpretation;
- the programmes are oriented toward training students for entry inro the profession;
- the courses are delivered by professional conference interpreters;
- teachers of interpretation have received teacher training related to interpretation;
- trainees' competence in conference interpretation is assessed in accordance with the professional entry requirements.

The latter point mentions competence in conference interpretation. The competence is defined as the system of knowledge which is needed to complete a given task, e.g. PACTE's[3] definition of translation competence: "the underlying system of knowledge, skills and attitudes needed to translate" (Hurtado Albir, 2017, p.295). A trend for demand-driven and competence-based

[2] 'AIIC' stands for International Association of Conference Interpreters.
[3] PACTE is the name of a research group from the Autonomous University of Barcelona. The name stands for Procés d'Adquisició de la Competència Traductora i Avaluació (see https://grupsderecerca.uab.cat/pacte/en).

education has clearly manifested itself in various domains of professional development over the last 20 years (e.g. Vandepitte, 2008). For example, the process that led to Common European Framework of Reference for language learning (CEFR) began in 1971 (Council of Europe, 2021). A similar competence framework for translators is being developed since 1997 (Hurtado Albir, 2017). Also, digital competence in education has been recognized as a priority at the European Commission (EC) level (European Commission, 2017). A set of competencies required in interpretation is also described in *The International Standard Classification of Education: Fields of Education and Training 2013* (UNESCO, 2015).

Competence approach is also being widely used elsewhere. In particular, the *Strategic plan for 2021 – 2025* by the Alliance for Health Policy and Systems Research, a World Health Organization (WHO) based partnership responsible for developing an agenda for rehabilitation research, acknowledges the important role of data-driven and experiential learning at WHO (World Health Organization, 2021). The *Rehabilitation Competency Framework* envisages the contribution of rehabilitation workers to education programmes at proficiency Level 2 and leading the curricula development at Level 4 (World Health Organization, 2020).

Demand-driven and competence-based approaches are evident in interpreter training and professional practices as well. Nowadays interpreters are expected to prepare for meetings using event-specific materials (Horváth, 2016). They are also expected to use a wide range of new technologies, ranging from social media (Poger, 2018) to digital pens (Orlando, 2016). One of the forms in which the demand-driven and competence-based approaches in interpreter training/ practice are implemented is *practice groups*, some of which have started to operate online during the pandemic (Fonseca, 2022). In order to meet the need for building and developing professional competencies of their members, such groups use a number of software solutions/ tools. These tools are intended for practitioners and trainees alike, and provide leverage for various important competencies, such as the accuracy of messaging, mastery of the target language, and the coherence of presentation (Deysel, 2018). The Translation[4] competencies are rooted in the widely recognized *Interpretive Theory of Translation* by Marianne Lederer and Danica Seleskovitch from ESIT Graduate School for Translation and Interpretation (ESIT)[5].

[4] This thesis explores the issues of interpretation (oral), however the term 'Translation' (with upper-case 'T') is used interchangeably to refer both to interpreting and translation as a cognitive process. This notation follows Gile (e.g. see Gile, 2009, p.263).

[5] ESIT stands for Ecole supérieure d'interprètes et de traducteurs http://www.univ-paris3.fr/bienvenue-sur-le-site-de-l-esit-63854.kjsp

## 2.2. Educational theories and approaches supporting professional practice

The theoretical mainstay of AIIC's practices is Interpretive Theory of Translation (by Marianne Lederer and Danica Seleskovitch from ESIT), complemented by functionalist approach (e.g. by Franz Pöchhacker from the University of Vienna), while the practical component is based on process-oriented approach by Daniel Gile (Université Paris 3 Sorbonne Nouvelle) (Baker, 2008). This is a combination of context-related, process-oriented and integrative approaches to interpreting. In other words, this is the educational framework that teaches trainees to integrate their skills and competencies in various ways depending on the interpreting task. The efficient adaptation to the task requires self-observation, structured thinking and awareness of various criteria. The Prototypes described in this thesis are built on the assumption that the target audience is at least superficially familiar with the principles described below.

### 2.2.1. Functionalist approach

Functionalist approaches to Translation postulate that "the linguistic form of the target text is determined by the purpose it is meant to fulfil" (Baker, 2008, p.115). If compared to learning theories, it seems to share some traits with behaviourism ("Learning Theories and Models - Summaries & Guides", n.d.). More specifically, functionalist theories argue that "the prospective function or purpose of the target text" determines the Translation process (Baker, 2008, p.116), while behaviourism seems to suggest that the learner behaviour is shaped and reinforced by the stimuli that come from the environment ("Learning Theories and Models - Summaries & Guides", n.d.). Thus, the purpose of the text (the expected effect of the Translation on the recipient) becomes a stimulus to a specific professional behaviour (e.g. the choice of Translation strategies). Franz Pöchhacker is an advocate of the functional approach and gives emphasis to communicative, context-related and product-based aspects of Translation. For example, while recognizing the invariably essential nature of such criteria as *accuracy* and *fidelity*, Pöchhacker builds on the assumption that "interpreting can and should be viewed within a conceptual spectrum from international to intra-social[6] spheres of interaction" (Pöchhacker, 2001, p.410). He also makes a case for more research into "(inter)personal qualities" that interpreters need in order to handle the likely "emotionally taxing" situations during the performance of the interpreting assignments and calls for greater regard to occupational profiles (Pöchhacker, 2014, p.148).

[6] See also *Socio-Translational Collaboration in Qualitative Inquiry: The Case of Expert Interviews* (Littig & Pöchhacker, 2014).

### 2.2.2. Interpretive Theory of Translation

The Interpretive Theory, formerly théorie du sens (Baker, 2008, p.53), was developed by Danica Seleskovitch of ESIT, Paris, in cooperation with Marianne Lederer and other colleagues. It describes Translation as a three-stage process: (a) the extraction of meaning from the source, (b) 'deverbalisation', which is 'forgetting' about the linguistic form of the source language, and (c) the re-expression of the meaning in the target language (Lederer, 2003).

As Gile (2009) puts it, there is observational evidence that conference interpreters tend to forget rapidly the exact form of the original utterance and replace it with their own mental representations ('packaging' in Gile's terminology), which takes a combination of linguistic and extralinguistic knowledge. This latter point about knowing how to act due to the integration of skills and abilities fits the definition of competence in Lasnier (as cited in Hurtado Albir, 2007)[7]. According to Hurtado Albir (2017), competence-based training (CBT) has its fundamentals in "cognitive-constructivist and socio-constructivist learning theories" (Hurtado Albir, 2017, p.14). Building a subjective representation during the deverbalization stage requires procedural knowledge, that is the know-how of approaching the Translation process (Beeby, 2009; Gambier, 2007) or "the ability to actually perform actions" (Gile, 2009, p.9). This knowledge is difficult to verbalise, it is processed automatically and is acquired through practice (Hurtado Albir, 2017). However, it can be taught and is being taught at programmes in Conference Interpreting, which means this is not only procedural knowledge of experts in Translation, but also the pedagogical knowledge that makes sense for learning (Lonka, 2018).

### 2.2.3. Process-oriented approach

Since its inception in the 1970s, Interpretive Theory of Translation (Baker, 2008) has become commonplace "in the form of a set of Translator training principles" (Gile, 2009, p.252). Although it has seen some criticism in various professional communities since 1990s, it still remains influential in training (Gile, 2009). According to Baker (2008), it continues to be the cornerstone of AIIC's best practices. One of the reasons is because this theory has offered a "practical, prescriptive and simple" approach to interpreter training (p.298).

[7] The original article by Lasnier is in French, which is not one of the author's working languages. For this reason, a reference from Hurtado Albir is provided.

At the heart of interpreter training curriculum, there are exercises (Gile, 2009). They mostly include Translations of texts from the source language into the target language followed by feedback from instructors. Gile is of the opinion that Translation exercises may need optimization, because they are "by definition *artificial*" (Gile, 2009, p.13), meaning that they are performed for pedagogical purposes rather than for making one's living working for a client, i.e. do not reproduce the real-life motivations and circumstances. The workaround he offers is to shift focus from the final version of the end product (i.e. the translated text) to *how* Translation happens and *why* errors occur (Gile, 2009). Gile calls it a process-oriented approach. He also advocates a strong pedagogical emphasis on the exercises that are directly related to interpreting techniques, such as note-taking, public speaking, and attention-sharing (Baker, 2008). Some of the useful training practices that Gile offers include:

- self-assessment of text comprehension by students;
- instructors asking 'leading' questions about specific barriers to understanding;
- classroom experiments with natural variability of texts;
- the Translation of texts that contain ambiguities;
- collective exercises aimed at "ad hoc Knowledge Acquisition", etc. (e.g. see pp.72, 98, 113, 150 in Gile, 2009).

This would be in line with metacognitive approach by Flavell ("Learning Theories and Models - Summaries & Guides", n.d.), and experiential education by Dewey ("A Visual Summary: 32 Learning Theories Every Teacher Should Know", n.d.). Namely, trainees are taught to monitor their knowledge and to focus on how they adapt it, which should eventually help them control their learning ("Learning Theories and Models - Summaries & Guides", n.d.). Interaction with the learning environment and guiding questions from teachers will likely provide some pedagogically meaningful experiences to learners ("A Visual Summary: 32 Learning Theories Every Teacher Should Know", n.d.). This demonstration of commonalities/ regularities would be an example of student-centered practice unlike teacher-centered convincing, lecturing or explaining. As Gile (2009) frames it, with process-oriented approach the teacher verifies the pathway that the learner has chosen and comments on potential problems rather than antagonizes students "by *imposing* one's own standards" and thus making learners reluctant to follow (p.15).

### 2.3. Development of Translation competences in the European region

This section gives a practical example of how the aforementioned principles have been embedded in a competence-based curriculum at an institution of higher education. As mentioned before, Translation training in Europe is by and large delivered via European Masters' Programme in Conference interpreting (EMCI).The overarching principles of EMCI are outlined in the Core Curriculum ("EMCI Core Curriculum", n.d.). Furthermore, this work will use the curriculum in consecutive interpreting (CI)[8] at Comillas university in Madrid as an example of Translation competence development in the European region. There are four reasons why this generalization would apply in the context of this thesis.

First, the Core Curriculum indicates that training in Simultaneous Interpretation (SI) skills "will build on the same kind as those used to practice consecutive interpretation" ("EMCI Core Curriculum", n.d., Section 3.4.), although with additional exercises aimed at developing SI specific skills[9].

Second, in the context of the functional approach and process-oriented approach (see sections 2.2.1. and 2.2.3. above), some of the competencies, such as accuracy and fidelity, may be considered invariant, at least for CI and SI.

Third, it is understood that the programme at Comillas university is recognized by the European Commission: Universidad Pontífica Comillas is an EMCI[10] member. The European Commission and the European Parliament cooperate with EMCI, as these postgraduate programmes comply with specific standards[11].

Fourth, the comparison of useful readings recommended by the EC and Comillas university show that the above theoretical framework is common and well renowned (see section *2.3.3.* Background readings recommended by the EC and Comillas university: some commonalities).

[8] Consecutive interpreting is defined as "a mode of interpreting in which the speaker makes a speech (or says a few sentences) whilst the interpreter takes notes. The interpreter then reproduces what the speaker has said for the audience" (European Commission, n.d., "Consecutive interpreting").
[9] Compare the description of CI and SI competences in the Core Curriculum: "At the end of the programme students shall be able to provide a fluent and effective consecutive interpretation of speeches reflecting professional conditions in conference settings, accurately reproducing the content of the original and using appropriate terminology and register" and "At the end of the programme students shall be able to provide a fluent and effective simultaneous interpretation of speeches of at least 20 minutes, accurately reproducing the content of the original and using appropriate terminology and register" ("EMCI Core Curriculum", n.d., Sections 3.3. – 3.4.).
[10] European Masters' Programme in Conference interpreting website is available at https://www.emcinterpreting.org/
[11] See for example https://ec.europa.eu/education/knowledge-centre-interpretation/conference-interpreting/conference-interpreting-explained_en

Therefore, it stands to reason that some of the characteristics of this programme and their practical implications be generalized to other top-notch programmes or their alumni.

### 2.3.1. Baseline competencies and the acquired competencies at the programme

This section describes the expected evolution of competence profiles at one of the EMCI institutions. As mentioned before, this curriculum is used as a 'representative specimen' of interpreting competencies. The description of these competencies is provided to explain the choice of the pre-defined criteria for the tool that the author proposes.

**Entry requirements at Comillas**

According to the description of the admission process ("Pruebas de admisión MUIC", 2021), candidates are expected, inter alia:

- to have the appropriate command (CEFR C2, C1) of their working languages;
- to be able to analyze and synthesize information in their working languages;
- to be familiar with international issues of the day;
- to be able to manage stress;
- to produce a speech in their own words and to demonstrate the appropriate level of reasoning.

They are NOT required:

- to have prior training in interpretation;
- to have prior training in translation;
- to focus on the structure of the speech that is provided for interpretation during the entry test.

**Acquired competencies**

At the end of the Master's degree programme the students are expected to develop the following competencies ("Competencias adquiridas", 2021), among other:

- General
  - Instrumental
    - *advanced ability to analyze and synthesize the information pertaining to conference interpretation;*
    - *to effectively manage the specific information pertaining to conference interpretation.*
  - Personal
    - *to think critically about the problems pertaining to their profession.*
  - Systemic
    - *to develop autonomous learning;*
    - *to have professional motivation for the quality of their job;*
    - *to apply their knowledge in practice.*
- Specific
    - *advanced knowledge of translation and interpreting techniques;*
    - *theoretical knowledge in the domain of conference interpretation.*

### 2.3.2. The competencies developed during the course in consecutive interpretation

The course in consecutive interpretation ("Guia docente 2021 – 2022. Ficha tecnica de la asignatura", 2021) aims to develop an advanced technical skill. It is also considered to be a foundation of skills in simultaneous interpretation. It is delivered in groups and relies on hands-on practical activities.

The course aims to develop the following competencies, among other:

- advanced ability to analyze and synthesize the information pertaining to conference interpretation;
- to effectively manage the specific information pertaining to conference interpretation;
- to think critically about the problems pertaining to their profession;
- to develop autonomous learning.

Furthermore, the course aims to develop the following techniques, among other:

- the usage of different types of memory;
- advanced oral presentation techniques;
- applying different coping strategies (processing speech);
- reproducing the structure, style and register of the speech in question;
- adapting to the working environment.

By comparing the acquired competencies with those that the course in consecutive interpretation (CI) aims to develop, one can appreciate that CI is the source of the fundamental and invariant competencies. Furthermore, one can appreciate that a clear distinction is made between the language learning competencies and Translation competencies (cf. "It [the Interpretive Theory of Translation] clearly marked the difference between interpreter training and language teaching" (Baker, 2008, p.298)). All of the above suggest that the Translation competencies, at least in the European region, is a well-shaped entity of its own which comprises a range of skills, abilities and know-hows needed to adapt to the task of rendering a Text (written/ oral/ signed) from one language into the other. The exact classification (or framework) of these competencies, however, is still under development, and although decoupled from language learning, "this framework should be similar to CEFR" (PACTE Group, 2018, p.112).

### 2.3.3. Background readings recommended by the EC and Comillas university: some commonalities

**Background reading: European Commission**

The Website of the European Commission's Knowledge Centre on Interpretation[12] lists a number of useful readings. These include, among other:

- *Basic concepts and models for interpreter and translator training* by **Daniel Gile**;
- *Pédagogie Raisonnée de l'Interprétation* by **Danica Seleskovitch** and **Marianne Lederer**;
- *Routledge Encyclopedia of Interpreting Studies* by **Franz Pöchhacker**.

[12] Available at https://ec.europa.eu/education/knowledge-centre-interpretation/conference-interpreting/research-and-background-reading/books-and-e-books_en

**Background reading: Comillas university**

The description of *Interpretation Theory* module at Comillas university refers to the following readings, among others ("Guia docente 2021 – 2022. Ficha tecnica de la asignatura", 2021):

- *Basic concepts and models for interpreter and translator training* by **Daniel Gile**;
- *Pédagogie Raisonnée de l'Interprétation* by **Danica Seleskovitch** and **Marianne Lederer**;
- *Quality Assessment in Conference and Community Interpreting* by **Franz Pöchhacker**.

All of the above seems to suggest that a number of European programmes in interpretation are based on the Interpretive Theory of Translation with modifications by ESIT and the University of Vienna. The general principles include competence-based learning, the usage of a competence framework, self-observation and peer-assessment, technical knowledge (*how to* knowledge) and background knowledge. For this reason, the author takes these principles into account in his Prototypes, and assumes a degree of familiarity with these principles among the target audience.

# 3. TOOLS TO SUPPORT COMPETENCE DEVELOPMENT FOR INTERPRETERS / TRANSLATORS

This section gives a brief outline of the commonly used tools for interpreter training/ practice and the commonalities that they have. Most of these tools are referred to in this thesis for two reasons. First, there is a large body of publications describing their positive impact in learning. Some of these publications are referenced in this chapter. Second, these tools have many times received positive reviews by members of a professional community of interpreters called *Techforword*[13], run by Joshua Goldsmith, a United Nations (UN) and European Union (EU) accredited interpreter and translator.

## 3.1. Tools used to support Translation practices

### 3.1.1. BabelNet

An artificial intelligence (AI) powered semantic network that provides semantic relations between concepts in many languages. It also has the functionality of encyclopedic dictionary. BabelNet allows users to choose concepts from over 40 domains, including linguistics, physics, medicine, education, business and many more. The system may be helpful in creating thematic glossaries for interpreting and/or language teaching, which saves a lot of time and effort. The tool is available at: https://babelnet.org/

### 3.1.2. OneClickTerms and N-Gram Phrase Extractor

Term extraction engines. They identify multi-word combinations and their frequencies; quickly process large amounts of texts and allow the export of results. They are useful for translators to identify term candidates in new texts. They can also come in very handy for language teaching, e.g. when there is a need to identify the most relevant new words/ word combinations. OneClickTerms has a free trial account, whereas N-Gram Phrase Extractor is a free online tool, and requires no registration. The tools are available at: https://terms.sketchengine.eu/ and https://www.lextutor.ca/phrases/n_gram/.

[13] Available at https://techforword.com/

### 3.1.3. Airtable

A flexible shared workspace, mostly based on Excel-type tables. Airtable works via a free account, as well as subscription. It can be helpful with terminology management: it allows users to add GIF files, videos, and pictures into the table. Thus, users can create very visual and organized glossaries. Airtable Supports sharing, and various interpreters have reported to be using Airtable for teamwork[14]. Furthermore, Airtable has a wide range of templates for various types of jobs, including project management. It supports co-authoring and commenting. Airtable also includes desktop, web, iOS and Android versions. The tool is available at: https://www.airtable.com/

### 3.1.4. InterpretBank

An AI powered desktop and web app for interpreters primarily designed to create, modify and manage glossaries, as well as to do a quick search while in the booth. It gives glossary suggestions (suggested translations); and collects relevant terms and their definitions at user's request from the Internet. InterpretBank relies, inter alia, on the IATE (Inter-Active Terminology for Europe[15]) database. Furthermore, it has a memorization functionality that helps users learn their glossaries. This tool is designed as a tool for interpreters, but also supports continuing professional development. The tool is available at: https://interpretbank.com/.

### 3.1.5. GoReact

An environment for teaching simultaneous and/or consecutive interpretation that fosters asynchronous learning, which is being used these days in many practice groups of interpreters. GoReact allows to upload the renditions (recorded translations) and synchronizes them with the source. It also allows to give feedback and synchronizes it with the timecodes in the recording. Users can upload their feedback in the form of texts, video or audio files. GoReact has various privacy settings, so it can be used in the classroom by multiple users. The tool is available at: https://get.goreact.com/.

[14] "There is, however, a confidentiality problem when using sensible customer data, but this could be solved by using encrypted solutions like interpretershelp.com or airtable.com" (Rütten, 2017, passage 16).
[15] Available at https://iate.europa.eu/home

As evident from the above descriptions, modern-day tools for interpreter training/ practice share a number of commonalities:

- extracting terms;
- creating thematic glossaries;
- building a network of relevant concepts;
- supporting teamwork;
- promoting feedback and interaction.

These commonalities have effect on competence development, as will be discussed below.

### 3.2. Impact of software tools on learning/ competence development

The publications on InterpretBank date back to 2009[16]. This tool has been widely used and studied since then. The founder of InterpretBank is Dr. Claudio Fantinuoli, a PhD in Applied Linguistics and former conference interpreter, currently Head of Innovation at KUDO Inc. and Lecturer at the University of Mainz[17]. He has extensively covered InterpretBank in his publications. According to Fantinuoli, interpreters are special learners whose task is to autonomously search for topic-related linguistic and extralinguistic material. The exposure to large amounts of data during this search makes their learning data-driven (DDL). In order to operationalize this learning, computer tools are helpful (Fantinuoli, 2017). The process of automated data extraction and glossary creation during DDL is called *Corpus-based terminological preparation for simultaneous interpreting* (Xu, 2018) or *Corpus-driven Interpreter Preparation* (CDIP) (Fantinuoli, 2017). Corpus-based approach has been recognized as the current mainstream in Translation studies, and was shown to promote subject-field understanding, the awareness of specialized schemata and precise phraseology (Nieto, 2014). Furthermore, the usage of corpus-based preparation has been shown to improve accuracy in simultaneous interpreting (SI) and the recall of terminology in the follow up (Xu, 2018). From the pedagogical point of view, InterpretBank has proven useful during the pilot study at the University of Bologna (Fantinuoli, 2016). The software was found to enhance the efficacy of terminology search even among the novice users and could be integrated into the simultaneous interpreter workflow in a user-friendly way (Fantinuoli, 2016).

[16] For example, *InterpretBank: Ein Tool zum Wissens-und Terminologiemanagement für Simultandolmetscher* (Fantinuoli, 2009).
[17] https://www.claudiofantinuoli.org/site/index.html

During a *Techforword* discussion about the feedback criteria (personal communication in the members only chat), a colleague interpreter said that they did not find word-level criticism helpful, because a professional interpreter is likely to be aware of the phrases or wordings that they had not used properly. What the colleague thought to be more helpful instead was patterns of delivery and content. This seems to be in line with the observation by Lonka, who suggests that senior students [i.e. more experienced learners] are likely to be a *reflective-collaborative* type of learners, that is, the most likely type to view learning as the construction of knowledge rather than the intake or use of knowledge. They tend to value reflection, metacognition[18] and collaborative learning (Lonka, 2021). It also adds up to the idea that with more experience in pursuing a goal, people become more willing to seek negative feedback rather than positive feedback (including, in language learning), because negative feedback increases "experts' sense that they were making insufficient progress" (Finkelstein & Fishbach, 2012, p.22). The role of self-reflection in stimulating responsible choices and behaviours of students is extensively described elsewhere (e.g. Newell, 2002).

All of the above underscores the complex nature of feedback: there seems to be no one-size-fits-all solution for giving feedback, and different categories of learners seem to benefit from different types of feedback. Smith and Harris (2014) make a case for tailoring feedback to the users' levels and commend GoReact as a useful tool for modifying learners' behaviours (and, consequently, performance) on specific tasks. The strengths of GoReact that the authors address specifically include: the availability of delayed (i.e. asynchronous) feedback, the cloud-based nature of the platform, the availability of peer-review, and time-coding of comments (i.e. the linkage of feedback to specific time points in the speech/ Translation).

BabelNet is mostly described in literature as a powerful multifunctional semantic network, mainly used for Natural Language Processing, Machine Learning and Machine Translation (Navigli, 2012). However, it is also a valuable resource for practitioners of Translation because it integrates the manual translations of Wikipedia pages into multiple languages complemented by WordNet resources (Navigli, 2021). This makes BabelNet a continuously-updated online multilingual encyclopaedic dictionary, which is obviously highly relevant for practical work in Translation.

The usage of language corpora for learning purposes has one more application, which is concordancing. Concordancing is defined as "a means of accessing a corpus of text to show how any given word or phrase in the text is used in the immediate contexts in which it appears"

[18] Metacognition is defined as "one's own cognitive processes and products or anything related to them, e.g. the learning-relevant properties of information or data" (Tarricone, 2011, p.2).

(Flowerdew, 1996, p.97). This approach has seen limited application for raising the language awareness of trainee translators (as cited in Gaskell, 2004, p.302), but has been more widely used in language teaching (e.g. Boulton, 2017; McEnery & Xiao, 2011). In a small survey, Lextutor has been shown to improve writing skills, although with a somewhat steep learning curve (Gaskell, 2004).

To summarize, the tools described in this section mostly promote autonomous learning, peer-learning and feedback, and information management. Arguably, they could be called hi-tech ecosystems rather than multifunctional tools.

# 4. PRESENTATION OF THE PORTFOLIO WORKS

The author of this work would like to propose a different type of tool than the ones described in the previous section. The iterations presented in this thesis are relatively lo-tech, visual, and customizable. They also have an offline functionality, which makes the proposed tool "portable" and cloud-independent. This combination of features should make one or all of the Prototypes a budget solution for expanding access to tools for collaborative and student-centered learning.

## 4.1. History of the prototype. Educational influences

The author was trained as a conference interpreter/translator in St. Petersburg State University, and later in St. Petersburg School of Conference Interpreting and Translation (SCIT), which is currently a member of the EMCI Consortium (European Masters in Conference Interpreting). Training at SCIT relies on pedagogical assistance by international organizations [e.g. see (Alexeeva & Antonova, 2011)], while some of the SCIT tutors are members of AIIC[19]. As mentioned in the previous chapter, an important part of the studies at EMCI schools is based on principles described by Daniel Gile and Marianne Lederer. Their theories seem to be in perfect agreement about the fundamental principles of interpreting/ translation and provide an efficient practical perspective of teaching students.

[19] For example, see: https://aiic.org/client/roster/clientRosterDetails.html?clientId=6178&clientRosterId=47&no_header=true&name=Ekaterina-SHUTOVA

In this section, the elements that are relevant for the proposed tool will be covered. According to Gile (2009), optimal theoretical components for interpreter training that are directly usable in teaching can be derived from the Interpretive Theory and functionalist theories. In author's opinion, interpreter training tools (and other practical components) should possess the following characteristics described by Gile (2009):

- **Being relevant to learner's needs**

They should directly deal with the problems and issues that learners face. Meanwhile they "should not contain many more concepts that can reasonably be considered *practically* useful to the trainees" (Gile, 2009, p.18). This means that learners should be provided with simple and user-friendly tools that resonate with their worldview, or at least address the same value system.

- **Being clear**

They should be modified so as to be less technical and more pedagogical. As a result of the modifications, the tools should teach trainees to take specific action in specific situations. This means that the tools should *not* be purely theoretical and 100% accurate in terms of the latest advances in translation studies. Instead, they should provide actionable (and somewhat simplified) algorithms. However, a common problem among language practitioners is the adoption of tools without proper testing of their operability (Jarvis & Daller, 2013), which is why practical, simple and validated tools are highly relevant.

- **Being taught after the proper introduction**

In Gile's practical experience, learners are not necessarily aware of some factors (for example, capacity limitations), and therefore are not always cognizant of important action to take (for example, not to take too long notes). An introductory explanation of the conceptual framework could be helpful. Another approach could be to create a scaffolded structure (such as the list of questions on the evaluation form) that sensitizes learners to specific issues.

- **Recalled repeatedly throughout the course**

The reference to the conceptual (or pedagogical for that matter) framework should regularly be made. It is especially true for feedback. As Gile puts it, "instead of telling him/her that the sentence makes no sense, it may be best to ask whether s/he has conducted a plausibility test" (Gile, 2009, p. 20).

Another principle that should determine some must-have features for the design of interpreter training tools is aligned to the three main stages of Translation[20] defined by Lederer (2003): "Understanding — Deverbalization — Re-expression". The explanation of each of the stages is provided below.

- Understanding a speech or a text "is a process in which sense is extracted from an aural or graphic sequence through the combination of cognitive inputs with linguistic meanings" (Lederer, 2003, p. 230).
- Deverbalization is "the phase that comes between the understanding of a text and its re-expression in another language. Verbal signs fade as cognitive and affective sense is grasped" (Lederer, 2003, p. 225).
- Re-expression is "finding a suitable formulation", or "a process based on the ideas understood, not on the words used to convey them" (Lederer, 2003, p. 35).

A simplified diagram of the main concepts of the Interpretive Theory that the author considers highly relevant in interpreter training is provided in the Appendix section.

The third element considered in the design of interpreter training tools is quality levels. Pöchhacker (2001) describes several levels of quality in interpreting which include:

- accurate rendition of the source;
- adequate expression in the target language;
- equivalent intended effect; and
- successful communicative interaction.

Many of these criteria have often been applied at various points throughout the author's training. For example, SCIT is an active proponent of the 'exposure interpreting' method, which means participating in conference simulations and exposing oneself to working in front of the real audience, followed by peer feedback and self-reflection (Alekseeva, 2011). This would be in line with systemic and needs-driven approach. The usage of self-assessment journals and peer-assessment sheets (Shutova, 2011) serves to sensitize students to various levels of Translation and structured approach to training.

[20] 'Translation' (with upper-case T) is used to collectively refer to interpreting and translation as a cognitive process (Gile, 2009).

Structured approach is widely used by various international language tests and certifications, such as DELE[21] examinations in Spanish, IELTS[22] and TOEFL[23] tests in English, or CELTA[24] course. These assessments are based on the competence model, and mostly corroborate with the Common European Framework of Reference for Languages (e.g. see Papageorgiou, 2015; or Instituto Cervantes, 2018). The fundamental entity of CEFR is *competencies*, defined as " the sum of knowledge, skills and characteristics that allow a person to perform actions" (Council of Europe, 2001). A research group from the Autonomous University of Barcelona (UAB) called PACTE is of the opinion that professional translation needs a competence framework that "should be similar to CEFR" in terms of its goals and approach (Hurtado Albir, 2018).

The relevance of competence approach in Translation is further testified by a number of standards: e.g. European Standard EN 15038: 2006 — Translation Services — Service Requirements[25] or US ASTM F2575-06 — Standard Guide for Quality Assurance in Translation. Both encompass the definitions of competences and make a case for various types of specific competences in translation. St. Petersburg School of Conference Interpreting and Translation (SCIT) works toward implementing a competence framework aligned with these standards and recognizes that "clear specification of the skills and competences that their graduates are expected to acquire" (Afonin, 2011) is an imperative for any educational institution.

All of the above has brought the author to conclude that interpreter training significantly benefits from linking theoretical explanations (although simplified as needed) with hands-on work. These links manifest themselves in implementing practical tools (such as journals and peer-evaluation sheets) during the 'exposure interpreting' sessions. The reliance on typical lesson frameworks [more widely known as 'lesson shapes' (Cambridge Assessment English, 2019)] is also helpful in delivering structured training.

As described in Section 3.1. Tools used to support Translation practices, modern-day practical tools offer technologically advanced ecosystems to promote information management, learner autonomy and peer-learning. They are mostly cloud-based and often operate a subscription model. However, in author's opinion, there seems to be a gap in functionality[26]. A cloud-independent, lo-tech, visual and customizable tool seems to be lacking (e.g. Walczyński, 2018). This sort of

[21] An official examination certifying the mastery of the Spanish language (https://www.dele.org/).
[22] International English Language Testing System (https://www.ielts.org/).
[23] Test of English as a Foreign Language (https://www.ets.org/toefl).
[24] Certificate of English Language Teaching to Adults (https://www.cambridgeenglish.org/teaching-english/teaching-qualifications/celta/).
[25] Superseded by ISO 17100:2015 since late 2015.
[26] Cf. "We see lots of ed-tech initiatives that target people who are already privileged. They've already got devices, they've already got good connectivity. Those are the people in the crosshairs of a lot of initiatives" (ELA, 2022, paragraph 15).

functionality would offer a budget solution for expanding access to tools for collaborative and student-centered learning. The Prototypes described below are different iterations of a *peer-assessment tool for interpreter training* that is aimed at bridging this gap.

## 4.2. History of the prototype. Early iterations and user experience

### 4.2.1. From numerical table to qualitative data

This section describes the search for compromise between assigning numerical values to an extensive list of performance indicators and conducting rapid assessments based on a few criteria and qualitative/semiquantitative ratings.

#### 4.2.1.1. 'Statistical' table

**Purpose**

The 'statistical' table is one of the earliest iterations of the tool presented in this thesis. It was designed in 2015 for private lessons of English. Between October and April 2015, the author gave private lessons of English to a group of 4 individuals (non-interpreters; non-linguists). The lessons covered various aspects of the language, such as grammar, vocabulary and syntax. The group requested to design a reporting form that would allow them to track their progress and see the areas to focus on. This way the tool shown in Figure 1 emerged. As a point of note, interpreter training has always been the mainstay of the author's pedagogical work. For this reason, the author has later adapted this tool to producing diagrams, such as one shown in Figure 6. This helped track the efficacy of vocabulary instruction of interpreters.

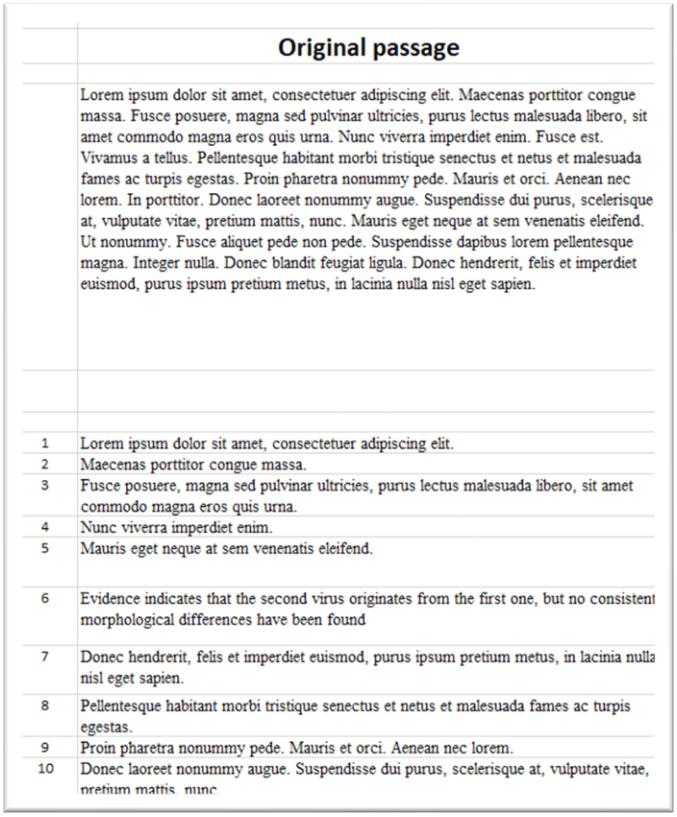
**Usage**

Figure 1 shows the overall appearance of the tool. Originally it served as the teacher's analytical tool. It was used to control the quality of retelling during the classes of speaking. This iteration has never been used for interpreter training, however, its modification (see *Radar chart* further) has seen practical implementation with private students, as well as at SCIT and St. Petersburg State University. One Excel sheet was used per each student, i.e. the tool was used for personalized monitoring.


| | | Words from to-learn list | | Grammar | | Variety of sentence models | Simple/non-simple sentence | Verbals |
|---|---|---|---|---|---|---|---|---|
| | | Repetition (+1) | First usage (+2) | Active voice (+1) | Passive voice (+2) | (1-3: +1/4-7: +2) | (+1/+2) | (+2) |
| | um dolor sit amet, consectetuer adipiscing elit. | | | | 2 | | 1 | 2 |
| 2 | as porttitor congue massa. | | HOST | 1 | | | 1 | 2 |
| 3 | Fusce posuere, magna sed pulvinar ultricies, purus lectus malesuada libero, sit amet commodo magna eros quis urna. | | OCCUR | 1 | | | 1 | |
| 4 | Nunc viverra imperdiet enim. | | | 1 | | | 2 | |
| 5 | Mauris eget neque at sem venenatis eleifend. | | | 1 | | | 2 | |
| 6 | Evidence indicates that the second virus originates from the first one, but no consistent morphological differences have been found | | INDICATE, EVIDENCE, ORIGINATE, CONSISTENT, FIND | | 2 | | 2 | |
| 7 | Donec hendrerit, felis et imperdiet euismod, purus ipsum pretium metus, in lacinia nulla nisl eget sapien. | FIND | DISTINCTIVE | 1 | | | 2 | |
| 8 | Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. | | | 1 | | | 2 | |
| 9 | Proin pharetra nonummy pede. Mauris et orci. Aenean nec lorem. | | IDENTIFY, DISTINGUISH | 1 | | | 2 | |
| 10 | Donec laoreet nonummy augue. Suspendisse dui purus, scelerisque at, vulputate vitae, pretium mattis, nunc. | | PENETRATE | 1 | | 1 | 1 | |
| | | 1 | 22 | 8 | 4 | 1 | 16 | 4 |

*FIGURE 1. Overall appearance of the 'Statistical' table tool.*

The “Original passage” field (A, B), in the left part of the tool (Figure 2), has a list of cells numbered from 1 to 10. These contain the learner-generated sentences. This form of note-taking allowed the teacher: (a) to keep track of the assignments that individual students completed, (b) to keep track of students’ individual progress.



| | Original passage |
|---|---|
| | Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Maecenas porttitor congue massa. Fusce posuere, magna sed pulvinar ultricies, purus lectus malesuada libero, sit amet commodo magna eros quis urna. Nunc viverra imperdiet enim. Fusce est. Vivamus a tellus. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Proin pharetra nonummy pede. Mauris et orci. Aenean nec lorem. In porttitor. Donec laoreet nonummy augue. Suspendisse dui purus, scelerisque at, vulputate vitae, pretium mattis, nunc. Mauris eget neque at sem venenatis eleifend. Ut nonummy. Fusce aliquet pede non pede. Suspendisse dapibus lorem pellentesque magna. Integer nulla. Donec blandit feugiat ligula. Donec hendrerit, felis et imperdiet euismod, purus ipsum pretium metus, in lacinia nulla nisl eget sapien. |
| 1 | Lorem ipsum dolor sit amet, consectetuer adipiscing elit. |
| 2 | Maecenas porttitor congue massa. |
| 3 | Fusce posuere, magna sed pulvinar ultricies, purus lectus malesuada libero, sit amet commodo magna eros quis urna. |
| 4 | Nunc viverra imperdiet enim. |
| 5 | Mauris eget neque at sem venenatis eleifend. |
| 6 | Evidence indicates that the second virus originates from the first one, but no consistent morphological differences have been found |
| 7 | Donec hendrerit, felis et imperdiet euismod, purus ipsum pretium metus, in lacinia nulla nisl eget sapien. |
| 8 | Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. |
| 9 | Proin pharetra nonummy pede. Mauris et orci. Aenean nec lorem. |
| 10 | Donec laoreet nonummy augue. Suspendisse dui purus, scelerisque at, vulputate vitae, pretium mattis, nunc |

*FIGURE 2. Magnified view of the 'Original passage' section.*

In the right part of the tool are the table (E) and diagrams (C, D, F) that Excel generates from the tabulated data. Sections of the table (see Figure 3) analyze the usage of words, grammar structures and syntax in each of the 10 sentences produced by learners.

For example, ‘Grammar’ section assigns different weights to the usage of active (+1 point) and passive (+2 points) voice. Each blue bar (Figure 4) shows one sentence in the active voice. Each red bar shows one sentence in the passive voice. Red bars are higher, which reflects their greater weight (more preferrable at this stage of learning). In other words, the diagram shows that the retelling consisted of 10 sentences, eight of which were in the active voice.

| Words from to-learn list | | Grammar | | Variety of sentence models | Simple/non-simple sentence | Verbals |
|---|---|---|---|---|---|---|
| Repetition (+1) | First usage (+2) | Active voice (+1) | Passive voice (+2) | (1-3: +1/4-7: +2) | (+1/+2) | (+2) |
| | | | 2 | | 1 | 2 |
| | HOST | 1 | | | 1 | 2 |
| | OCCUR | 1 | | | 1 | |
| | | 1 | | | 2 | |
| | | 1 | | | 2 | |
| | INDICATE, EVIDENCE, ORIGINATE, CONSISTENT, FIND | | 2 | | 2 | |
| FIND | DISTINCTIVE | 1 | | | 2 | |
| | | 1 | | | 2 | |
| | IDENTIFY, DISTINGUISH | 1 | | | 2 | |
| | PENETRATE | 1 | | 1 | 1 | |
| 1 | 22 | 8 | 4 | 1 | 16 | 4 |

*FIGURE 3. Assessment table.*

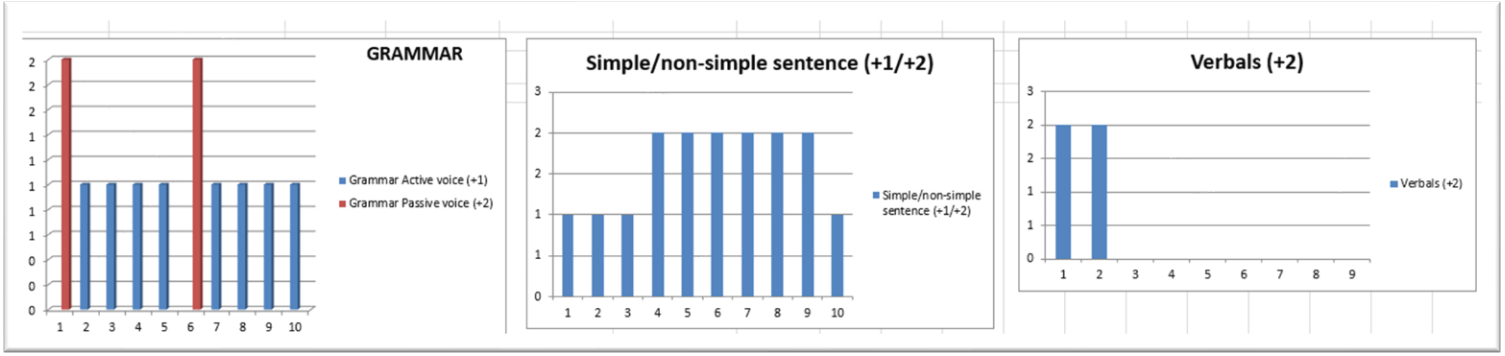



*FIGURE 4. Diagrams derived from the tabulated data.*

Section called ‘Words from to-learn list’ has two columns. The left column (‘Repetition’) is used to take note of the ‘favourite’ words. E.g. the original contains such words as ‘detect’, ‘reveal’, or ‘identify’. However, the student preferred to use ‘find’ in all instances instead. The teacher entered ‘find’ in this column in order to encourage this student to use other synonyms of ‘find’ in future. Conversely, the right column (‘First usage’) takes special note of the student’s using the words from the to-learn list. There is no diagram for this table because with many words in the to-learn list it would be too busy with information.

Likewise, columns ‘Simple/ non-simple sentence’ and ‘Verbals’ and the respective diagrams represent the usage of simple or compound/ complex sentences and infinitives/ participles/ gerund. One can appreciate that the student mostly used non-simple sentences (six high blue bars out of ten). However, they used verbals only on two occasions.

To sum up, all the data shows that this student was enthusiastic about the usage of new words and non-simple sentences (which happened to be compound in this case, i.e. linked with *and* or *but*), however, they relied almost exclusively on the active voice and did not seem to give justice to verbals. This warranted more emphasis on verbals and more attention to complex sentences (linked with *although*, *while*, *because*, *as soon as*, etc.).

**Findings**

Although informative, this tool has proven difficult for synchronous routine monitoring. A more traditional way of simply taking notes with pen and paper seemed to work better. Jotting down the most noteworthy pieces (such as awkward wordings, inaccurate phrases, or, on the contrary, examples of efficient use of language) may seem unstructured and imprecise, but it has proven more handy than filling in the Excel table.

When offered to students as a form of self-monitoring/ peer-observation, the tool was modest success. A brief period of excitement with diagrams, numbers and measurements was followed by the inevitable questions about their comparability, reliability and meaning. This has brought the author to conclude that his private practice should probably make use of qualitative rather than quantitative tools. On the other hand, structured approach and visual representation remained high on the author’s agenda.

### 4.2.2. From detailed criteria to broad wordings

This section shows the transition from excessively detailed assessment criteria to less stringent and broader descriptions.

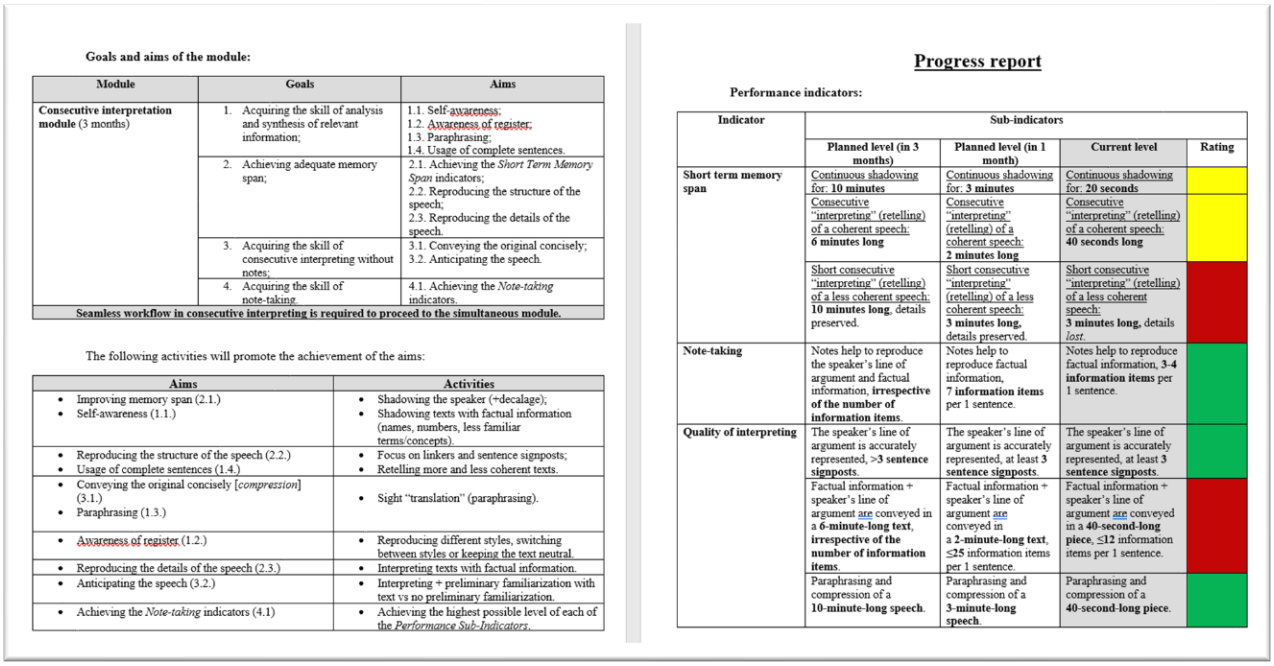
#### 4.2.2.1. Progress report

**Purpose**

This iteration of the tool was designed at the end of 2015 for a group of private students of interpreting. In order to promote cooperative learning, the author introduced an electronic template for interim reporting.

**Usage**

The report form is a two-page spread (see Figure 5). The page on the left outlines the goals, aims and activities. The page on the right shows progress toward the goals. Colour codes on the right characterize the achievement of the desired performance by the end of the reporting period (Green – achieved, Yellow – almost there, Red – not achieved). The sheet on Figure 5 shows progress two weeks into the course in consecutive interpreting. One can appreciate that the progress of the student (Figure 5) was mixed, with three activities progressing as planned, two activities requiring more effort, and two activities being in the red zone (i.e. the student underperforming).



**Goals and aims of the module:**

| Module | Goals | Aims |
|---|---|---|
| **Consecutive interpretation module** (3 months) | 1. Acquiring the skill of analysis and synthesis of relevant information; | 1.1. Self-awareness;<br>1.2. Awareness of register;<br>1.3. Paraphrasing;<br>1.4. Usage of complete sentences. |
| | 2. Achieving adequate memory span; | 2.1. Achieving the *Short Term Memory Span* indicators;<br>2.2. Reproducing the structure of the speech;<br>2.3. Reproducing the details of the speech. |
| | 3. Acquiring the skill of consecutive interpreting without notes; | 3.1. Conveying the original concisely;<br>3.2. Anticipating the speech. |
| | 4. Acquiring the skill of note-taking | 4.1. Achieving the *Note-taking* indicators. |
| **Seamless workflow in consecutive interpreting is required to proceed to the simultaneous module.** | | |

The following activities will promote the achievement of the aims:

| Aims | Activities |
|---|---|
| • Improving memory span (2.1.)<br>• Self-awareness (1.1.) | • Shadowing the speaker (÷decalage);<br>• Shadowing texts with factual information (names, numbers, less familiar terms/concepts). |
| • Reproducing the structure of the speech (2.2.)<br>• Usage of complete sentences (1.4.) | • Focus on linkers and sentence signposts;<br>• Retelling more and less coherent texts. |
| • Conveying the original concisely [*compression*] (3.1.)<br>• Paraphrasing (1.3.) | • Sight "translation" (paraphrasing). |
| • Awareness of register (1.2.) | • Reproducing different styles, switching between styles or keeping the text neutral. |
| • Reproducing the details of the speech (2.3.) | • Interpreting texts with factual information. |
| • Anticipating the speech (3.2.) | • Interpreting + preliminary familiarization with text vs no preliminary familiarization. |
| • Achieving the *Note-taking* indicators (4.1) | • Achieving the highest possible level of each of the *Performance Sub-Indicators*. |

**Progress report**

**Performance indicators:**

| Indicator | Sub-indicators | | | |
|---|---|---|---|---|
| | **Planned level (in 3 months)** | **Planned level (in 1 month)** | **Current level** | **Rating** |
| **Short term memory span** | Continuous shadowing for: **10 minutes** | Continuous shadowing for: **3 minutes** | Continuous shadowing for: **20 seconds** | |
| | Consecutive "interpreting" (retelling) of a coherent speech: **6 minutes long** | Consecutive "interpreting" (retelling) of a coherent speech: **2 minutes long** | Consecutive "interpreting" (retelling) of a coherent speech: **40 seconds long** | |
| | Short consecutive "interpreting" (retelling) of a less coherent speech: **10 minutes long**, details preserved. | Short consecutive "interpreting" (retelling) of a less coherent speech: **3 minutes long,** details preserved. | Short consecutive "interpreting" (retelling) of a less coherent speech: **3 minutes long,** details *lost*. | |
| **Note-taking** | Notes help to reproduce the speaker's line of argument and factual information, **irrespective of the number of information items**. | Notes help to reproduce factual information, **7 information items** per 1 sentence. | Notes help to reproduce factual information, **3-4 information items** per 1 sentence. | |
| **Quality of interpreting** | The speaker's line of argument is accurately represented, **>3 sentence signposts**. | The speaker's line of argument is accurately represented, at least **3 sentence signposts**. | The speaker's line of argument is accurately represented, at least **3 sentence signposts**. | |
| | Factual information + speaker's line of argument are conveyed in a **6-minute-long text**, **irrespective of the number of information items**. | Factual information + speaker's line of argument are conveyed in a **2-minute-long text**, **≤25** information items per 1 sentence. | Factual information + speaker's line of argument are conveyed in a **40-second-long piece**, **≤12** information items per 1 sentence. | |
| | Paraphrasing and compression of a **10-minute-long speech**. | Paraphrasing and compression of a **3-minute-long speech**. | Paraphrasing and compression of a **40-second-long piece**. | |

*FIGURE 5. Overall appearance of the Progress report tool.*

**Findings**

Although devoid of numbers, weights and diagrams, this tool still received mixed feedback. Excessive text, the list of ten aims (from 1.1. to 4.1.) and nine activities were reported to take up too much time and concentration for a routine self-monitoring/ peer-assessment tool. Furthermore, the non-binary nature of colour codes (pass – fail – neither) was not sensitive enough for monitoring progress over time or making comparisons. Furthermore, in order to assign the definitive colour code, inputs from several students were averaged. The testing accuracy of this procedure is, of course, open to discussion. However, the tool was deemed acceptable for interim one-to-one reporting, specifically because of promoting structured assessment. As students put it, "breaking the performance down into components" made their learning more focused.

This experience has made the author to conclude that a better tool should provide:

a) breakdown into criteria;
b) reasonably sensitive rating of these criteria;
c) relatively simple assessment procedure; and
d) understandable visual representation of progress.

These inputs have paved the way for the further development of tools by the author.

### 4.2.3. Current prototypes. Prototype 1

#### 4.2.3.1. Radar chart, paper version

**Purpose**

Prototype 1 was designed in 2017 in order to meet the need for a simple and visual tool for giving structured feedback in consecutive and simultaneous interpretation. The abovementioned need for breakdown into criteria, simple and sensitive assessment, and understandable visual representation of progress have led to the creation of Prototype 1.

**Usage**

This tool was used in four contexts:

- *Annual workshops/webinars in medical interpreting at SCIT (annual 1-2 day events, held regularly between 2015 and 2020; the tool in use since the end of 2016).*

Annual workshops/ webinars were 1 or 2-day events for a group of post-degree trainees of interpretation. The author conducted these meetings within the framework of the *alumni pedagogical assistance*. The overarching aim of these workshops was to familiarize the students with the core (invariant) skills and competencies in interpretation that helped handle the job at medical conferences. The groups were usually offered thematic assignments, and the author was expected to give actionable structured feedback. In plain terms, students wanted to get the feel of competencies required to succeed in medical interpreting. They also wanted a sort of roadmap that they could relate to their current progress. In author's experience, efficient feedback is often a matter of trust: students can find it difficult to internalize feedback from someone whom they have met only on a few occasions before. It was therefore important to make the assessment as objective as possible. In particular, through peer feedback. Thus, the tool served to incorporate multiple inputs from the whole group and present it in a visual and concise form.

- *CoreReview events at St. Petersburg State University (two events in 2019).*
  The author designed CoreReviews as a follow up activity for the alumni of the post-degree programme in interpreting. The alumni attended these one-day events together with the current students. The events were strictly thematic, which means they addressed a specific technique, such as the translation of numbers, or presenting statistics. While the alumni had the opportunity to refresh the existing skills, the current students could build up new ones. Therefore, these events were a form of peer learning. The tool served to bring together multiple inputs from a heterogenous ad-hoc group of learners and produce structured feedback. The role of this feedback was to provide the participants with a roadmap indicating their strengths and weaknesses.

- *For post-degree training in consecutive interpretation at St. Petersburg University (twice a month between October and December 2019, i.e. on 6 occasions).*
  The tool was used for the point assessment of performance every two weeks. It provided structured assessment based on a pre-defined set of quantitatively measured criteria. The usage of standard forms over the time has also contributed to tracking progress and adjusting the learning routines accordingly.

- *For in-house training in consecutive interpretation (2 staff members of a private company).*
  Between October and December 2016 the author conducted regular training sessions in

interpreting for two technical staff members at a trade company. Because the scope of these persons' activities included ad-hoc translation/ interpreting, there was a need for actionable and structured feedback on their work, as well as a roadmap. This tool provided numerical assessment on a scale of 0 to 5 and included three groups of criteria: verbalization, short-term memory, and delivery, which is in line with Daniel Gile's Tightrope Hypothesis[27] (e.g. see Gile, 1999; Gile, 2009; or Baker, 2008). Verbalization acts as a proxy of listening (because listening comprehension cannot be immediately assessed when an interpreter is making their rendition), and delivery (the public speaking aspect of interpreting) acts as a proxy of the target text production.

The paper versions had been in use until 2020. They were distributed to trainees at the beginning of the lesson and then collected during the break, and scores were entered into the Excel template (Figure 7). Diagrams (charts) were then created and shown on the teacher's PC or e-mailed to individual interpreters/ trainees and discussed after the break. Polygons on radar charts represented the scores assigned for each of the skills, which allowed interpreters to see at a glance what impression their work has produced. During the COVID pandemic in 2020, the first transition to the fully electronic format took place. Each of the students was sent a blank template which they filled in and returned by e-mail. The data entry was still happening manually during the break, but this time it was quicker and easier. Individual results were sent either as Excel files or screenshots depending on the device that students were using (smartphone users asked for a picture, as Excel would not display diagrams on some of the phones).

[27] The Hypothesis which, according to Gile (1999), has massive anecdotal reports, postulates that interpreters simultaneously perform multiple cognitive operations in order to complete three parallel tasks: (1) to listen and to analyze the original text; (2) to produce the target text and to self-monitor the quality of one's own work; (3) to engage short-term memory. These tasks draw on interpreter's total available resources and may lead to a *Tightrope situation*, i.e. working close to the maximum capacity.

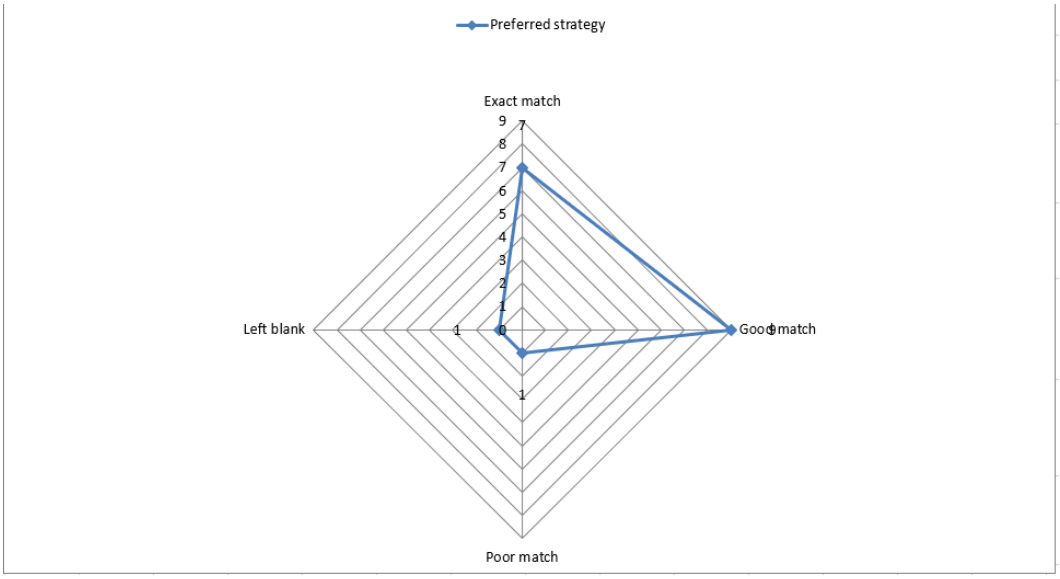



*FIGURE 6. The usage of radar charts for visualizing the vocabulary strategies of interpreting students (gap filling exercise).*

Furthermore, radar charts were also found to provide more insight into vocabulary strategies of individual students (Figure 6). This has become an efficient adaptation of the earlier iteration (Figure 1). Charts helped visualize the strategies that interpreters/ trainees used when asked to do the *gap fill* and *lexical neighbours* exercises. Either exercise had gaps that ideally had to be filled with previously learned words or word combinations. There were four scenarios:

a) to leave the gap blank;
b) to fill the gap with the exact match (previously learned word or the combination of words);
c) to fill the gap with a good match which is similar to the exact match in meaning and form;
d) to fill the gap with a poor match.

Interpreters were expected to avoid scenario (a) and to consider scenarios (b) and/or (c). If the choice of strategies was projected to a radar chart, the resulting polygon would be expected to show up in the top right corner of the chart (see Figure 6). The tool was published at *Zenodo* repository (https://zenodo.org/record/3755784#.YlHaC8hBxM0).

**Findings**

This table has proven to be very practical, which allowed its usage in various modifications. Visual charts significantly facilitated discussion about the interpreter performance and simplified progress tracking. The tool has also allowed to elicit the vocabulary strategies of interpreters and guide interpreters throughout their learning. The author believes that the usage of this tool is

capable of stimulating more actionable peer feedback. In his pedagogical work, the author has observed intrinsic barriers to talking about the gaps in performance as well as certain difficulty giving structured feedback. In either case, the resulting feedback is difficult to act upon.

On a scale of 0 to 5, with 1 being poor and 5 being excellent, how would you rate the interpreter performance?

| Interpreter | Verbalisation | | | | Short-term memory | | Delivery | |
|---|---|---|---|---|---|---|---|---|
| | Interpreter is articulate | Register is appropriate | Language interference (calque) is avoided | All sentences are complete | The speaker's line of argument is accurately represented | Important factual information (names, numbers) is conveyed | Interpreter engages the audience | The original meaning is conveyed in a concise way (compression) |
| 1 | | | | | | | | |
| 2 | | | | | | | | |
| 3 | | | | | | | | |
| 4 | | | | | | | | |
| 5 | | | | | | | | |
| 6 | | | | | | | | |
| 7 | | | | | | | | |
| 8 | | | | | | | | |

*FIGURE 7. A handout for peer-assessment. It was later replaced by the electronic version.*

For example, in some contexts the participants actually met each other for the first time. In such cases, the participants usually showed some reservation about giving negative feedback, and mostly provided positive feedback framed in somewhat general wordings. In other situations, for example during the one-off workshops, the lack of scaffolding (such as evaluation sheets) complicated feedback. The learners' routine feedback procedure sometimes used to be different from what the author had to offer. For example, the author focused on a few specific aspects of interpreting that usually received less detailed attention. In either case, the introduction of a template/ form facilitated clearer communication and more comprehensive feedback.

The concise and lo-tech nature of this tool (table and vocabulary charts) promoted its practical implementation during the private classes of interpreting as well as workshops with interpreting trainees at SCIT. It was also used for post-degree training at St. Petersburg State University between 2017 and 2019.

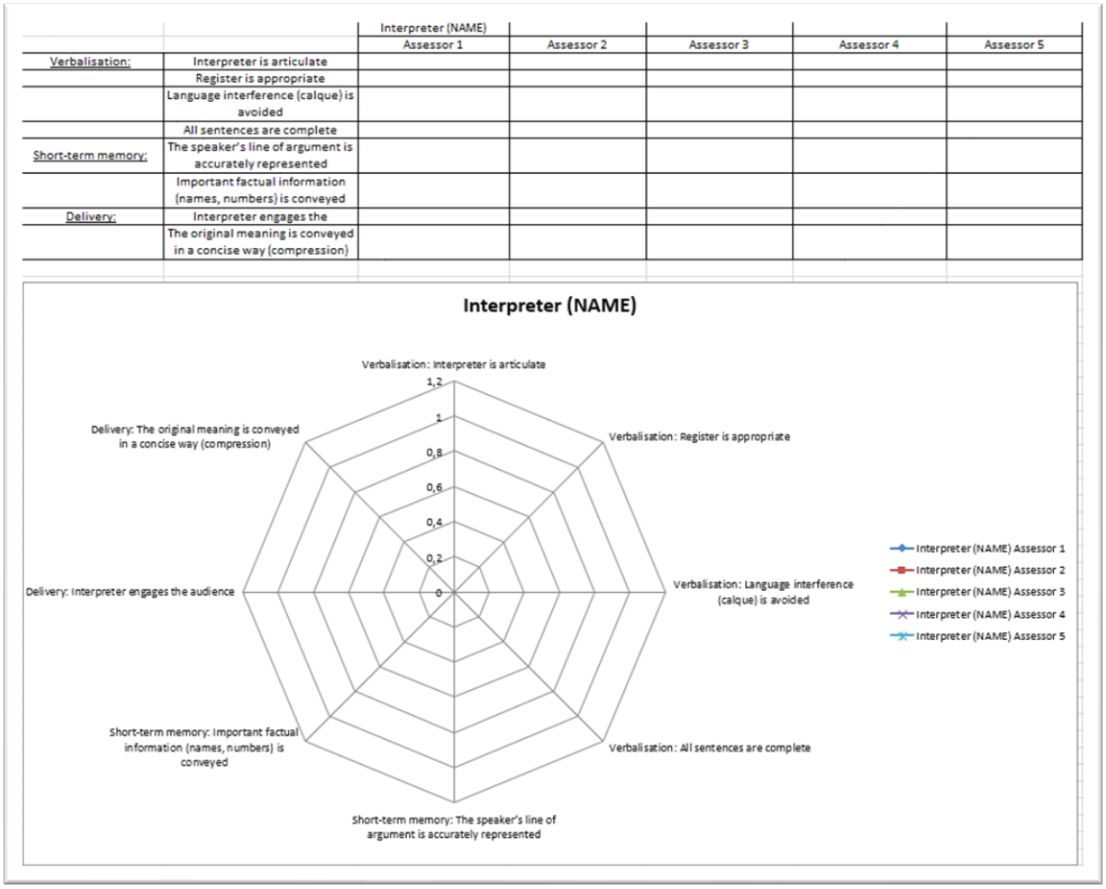



*FIGURE 8. Teacher's version of the Excel template.*

### 4.2.4. Current prototypes: the evolution at Oulu. Prototypes 2 and 3

This section describes Prototype 2 and Prototype 3. These are two consecutive iterations of the tool that have different applications, implementations and functionality. Prototype 2 emerged out of cooperation with peers at Oulu University of Applied Sciences (OAMK). Prototype 3 was created at request of the author by Kenny Tran, a freelance full-stack developer. Both iterations were created between September 2021 and February 2022, during the author's studies at OAMK.

#### 4.2.4.1. Prototype 2: MS Forms questionnaire

**Purpose**

Prototype 2 (Figure 10; Figure 11) has evolved as a tool to meet the need for a simple and visual tool for giving structured feedback in a wider range of contexts. The need for Prototype 2 arose because of some downsides of its forerunner, Prototype 1. As mentioned before, Prototype 1 facilitated visual feedback based on multiple inputs. It also promoted more comprehensive and

open exchange of opinions. On the other hand, the administration of the paper form was time-consuming. The author usually collected the paper evaluation sheets before the break, transferred the data into the Excel form manually during the break and presented the results after the break. Furthermore, this was not sufficiently interactive.

In 2020, the author's webinar at SCIT was delivered online for the first time because of the COVID pandemic. The administration of paper forms was not feasible, which is why the author sent electronic evaluation sheets. The process was equally time-consuming, because the author collected the electronic forms via e-mail during the break. However, this was the first step toward the fully electronic version of the feedback tool.

The pending issues were:

- to make the assessment criteria self-explanatory;
- to speed up the administration of the assessment (less burden for the tutor and greater interactivity for the students);
- to make the output more meaningful.

All of the above obviously required further updates to Prototype 1. The author continued his work on the tools for interpreter training at OAMK. Inputs from courses in *Service Design* and *Education Theory* as well as discussions with peers and tutors have brought the author to conclude that the scope of Prototype 1 could probably be extended to incorporate language teaching.

**Usage**

In view of the extension to language learning, Prototype 2 is meant to facilitate (a) interpreter training in EMCI institutions; (b) continuing professional development in practice groups of experienced interpreters; (c) structured feedback among the learners of foreign languages (Figure 9; see section 5.5. MEE workshop results for a more detailed explanation). It has not been used in practice so far: it has only been tested with several groups of audiences (MEE peers; focus group). Brief overview of these findings is provided in this section, whereas more thorough analysis is provided in sections 5.4. and 5.5.

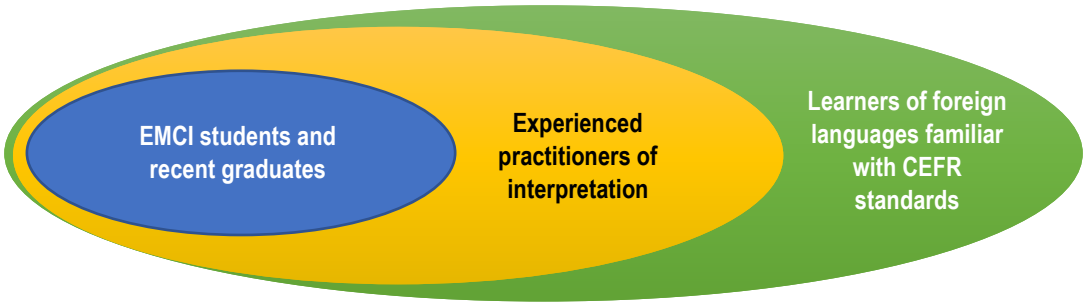



*FIGURE 9. Core and extended-criteria target audiences.*

Focus group findings:

Prototype 2 was presented to a Focus group (FG) of interpreters in April 2022. Based on the feedback, the positive findings are:

(a) the FG members found Prototype 2 to be relatively easy to handle;

(b) the FG members found various pre-defined criteria to be relevant for their assessment;

(c) the FG members were mostly comfortable with the numerical ratings.

The negative findings are:

(a) some FG members preferred alternatives to numerical ratings;

(b) the FG members lacked a free-text field for feedback;

(c) many of the FG members were particularly conscious about the wordings of the criteria;

(d) the Prototype did not allow for the modulatory effects of different factors (which could be solved, for example, by introducing relative weights).

MEE findings:

Based on the feedback from MEE peers, the positive findings are:

(a) the participants thought that Prototype 2 promoted more open and comprehensive feedback;

(b) the participants believed that Prototype 2 promoted systematic assessment;

(c) the participants believed that Prototype 2 was user-friendly and easy to handle.

The negative findings are:

(a) the participants did not consider the numerical ratings to be optimal;

(b) the participants obviously lacked a free-text field for feedback;

(c) the participants did not find all the criteria to be self-explanatory.

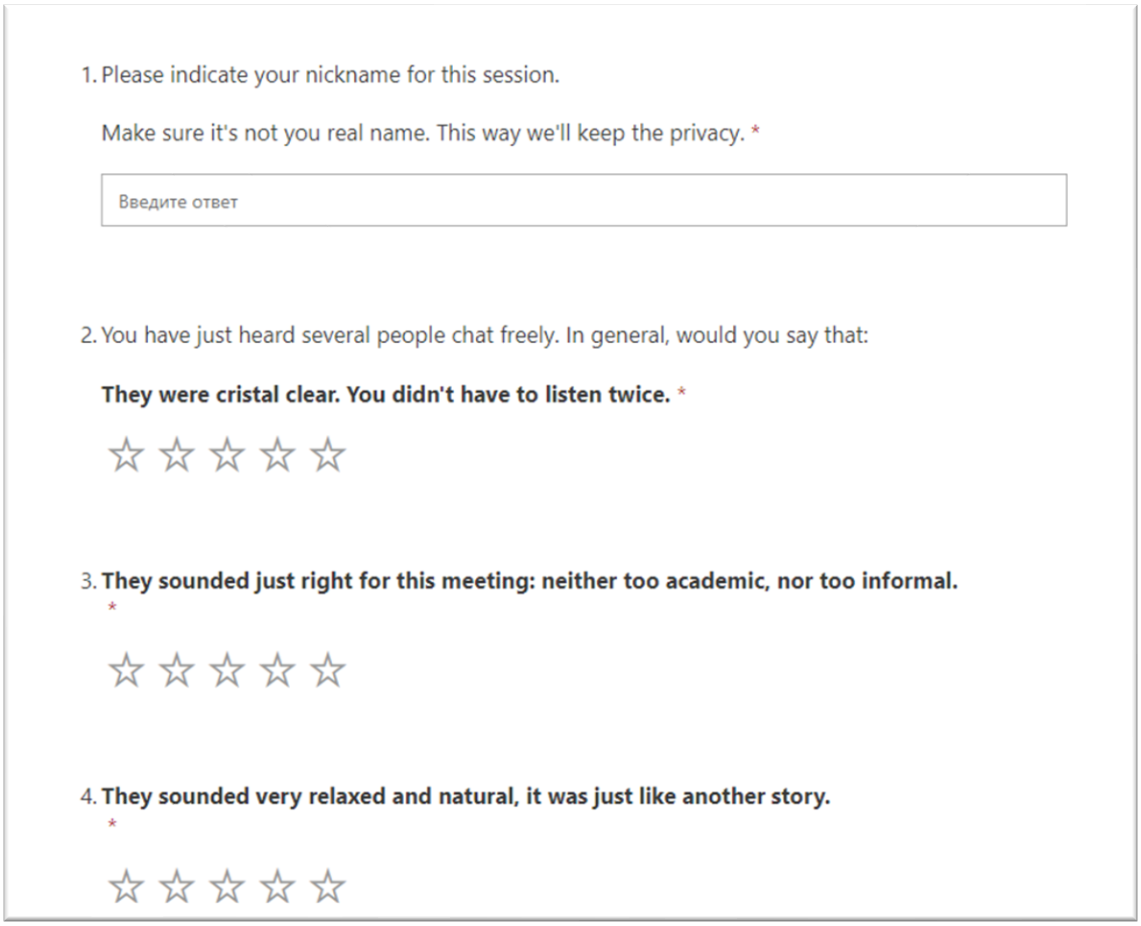



*FIGURE 10. Prototype 2. The UI part. Users can solely grade performance based on a pre-defined set of questions. All questions are marked as mandatory.*

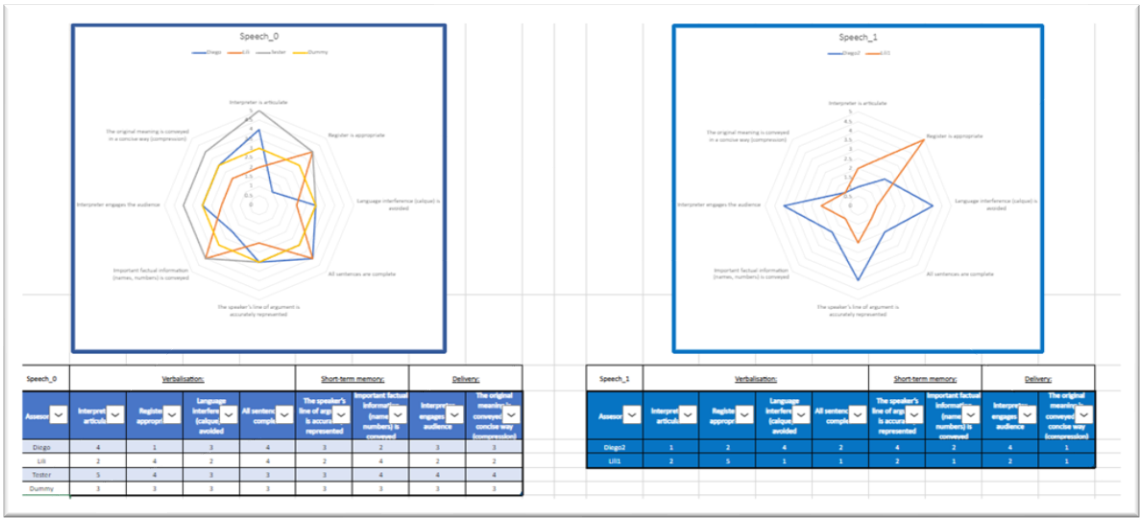



*FIGURE 11. Prototype 2. The Excel tables and radar charts. Data from MS Forms is transferred to Excel by means of MS Power Automate. The users do not have access to this Excel. Thus, the resulting chart had to be sent to users by the author after the poll.*

#### 4.2.4.2. Prototype 3: TailwindUI tool

**Purpose**

Prototype 3 is a web application that serves the same purpose as Prototype 2, i.e. to meet the need for a simple and visual tool for giving structured feedback in a wider range of contexts. However, as mentioned before, it has different applications, implementation and functionality. Prototype 2 lacked free text fields; not all the criteria were clear/ acceptable for the target audience; the usage of Prototype 2 was impossible without the moderator (someone who has access to final results and forwards these to the participants). Prototype 3 seeks to bridge this gap. It provides a user-friendly user-managed form that can be accessed and used autonomously.

**Usage**

Prototype 3 is the app with TailwindUI[28] interface. It is uploaded to Neltify server and is made available from mobile devices and desktops via an ordinary web-link. It allows to freely add any number of categories. The app has sliders that allow users to rate performance on a scale of 0 to 5. Radar charts are created automatically. The focus group meeting took place online, which is why the online version of Prototype 3 was provided. However, it also has an offline version that can be run on a local computer/ inside a LAN of an educational institution.

[28] See https://tailwindui.com/.

The prototype has a simplified form of personal areas: the access to grading is password-protected, and the results are made available only to the person being assessed. Thus, the prototype is designed to provide confidential grading with the results instantly available. No admin is required, and users can access the tool on their own. The prototype is not designed to collect any sensitive information, nor does it require registration, sign-in with any social media, etc. It has a simple authentication functionality. At the first entry every user receives an automatically generated username. The user then creates a password which – in combination with the username – works as a key to access the system. The session (called an assessment) is recorded as a separate file and also receives an autogenerated name. These names are deliberately made meaningless, and essentially they are a combination of random symbols. Prototype 3 has not seen practical usage so far. It has only been tested with (a) the focus group (FG) of interpreters in April 2022 for the purposes of this study; and (b) a colleague interpreter as a matter of personal communication and sharing information about the author's pedagogical approach.

**Findings**

When presented to the focus group of professional interpreters (see section 5.2. for the characteristics of the group), Prototype 3 revealed the following:

- Professional interpreters tend to draw a distinction between 'general' and 'technical' skills/ competencies. For this audience, the numerical rating of skills/ competencies that Prototype 3 suggests would be relevant solely in the context of formal assessment.

- The 'general' and the 'technical' skills/ competencies have very few overlaps. Rather, the 'technical' skills/ competencies may be viewed as the practical manifestation of the 'general' skills/ competencies.

- Prototype 3 would be more relevant for the 'technical' assessment.

- The assessment procedure is context-dependent and should take into account some variables, such as the type of Translation and the applicability of the assessment criteria. Furthermore, professional interpreters seem to object against any scaffolding, such as pre-defined lists of criteria and prompts. The degree of flexibility provided by Prototype 3 (no pre-defined categories) seems to be an advantage.

- The focus group seems to be in favour of preliminary familiarization with the tool. Although relatively straightforward, Prototype 3 would benefit from a preliminary explanation about the way it works.

Possible advantages of Prototype 3 may include: (a) fully autonomous operation and reporting; (b) anonymity/ privacy; (c) visual representation of results; (d) free choice of criteria. Possible disadvantages may include: (a) lack of introductory explanations or prompts; (b) somewhat complicated architecture of the current implementation (need for Next.js framework, reliance on a database file).

Detailed analysis of the FG findings is provided in section 5.4. Focus group results. Further avenues of Prototype 3 development and its practical usability are presented in the Conclusions and reflection section.

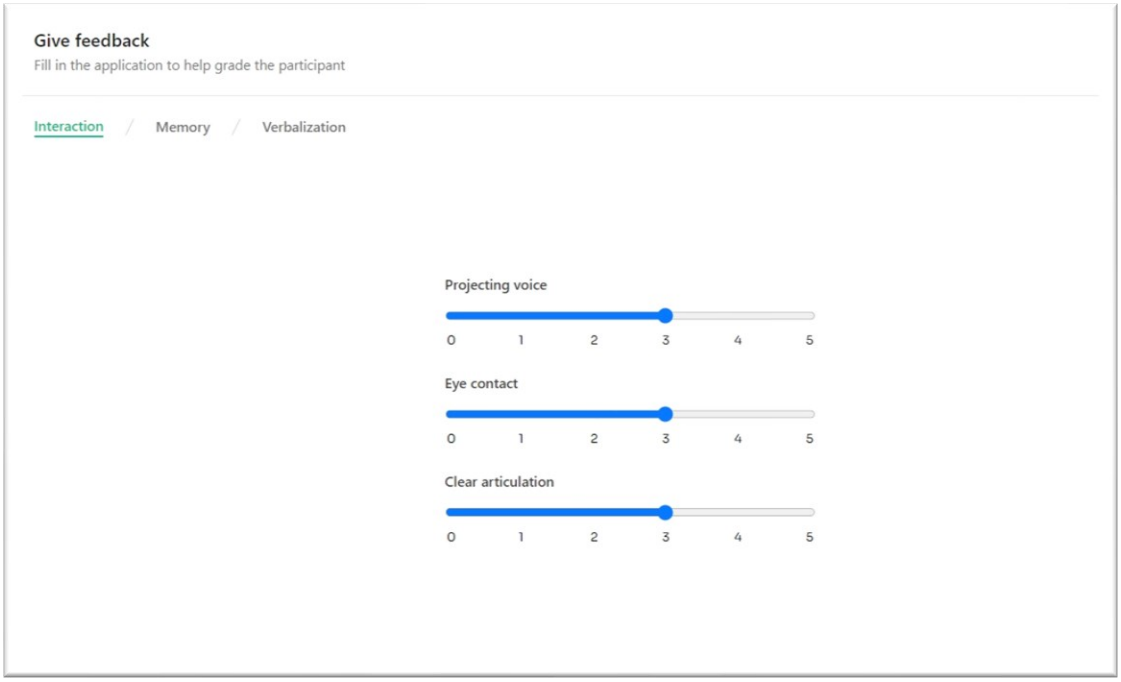



*FIGURE 12. Prototype 3. Sliders that provide a scale from 0 to 5. Users define how many categories will be in the assessment and what their names are.*

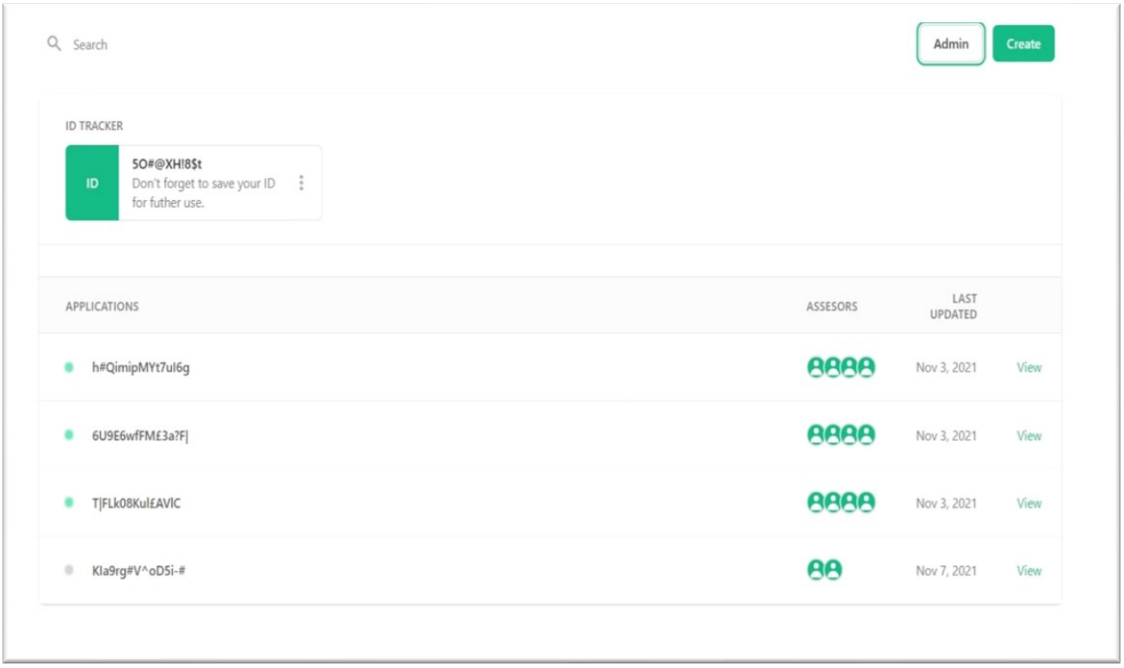



*FIGURE 13.Prototype 3. TailwindUI interface. The admin panel. Users have similar panels. For the sake of privacy, the IDs represent a random set of symbols, rather than the real names.*

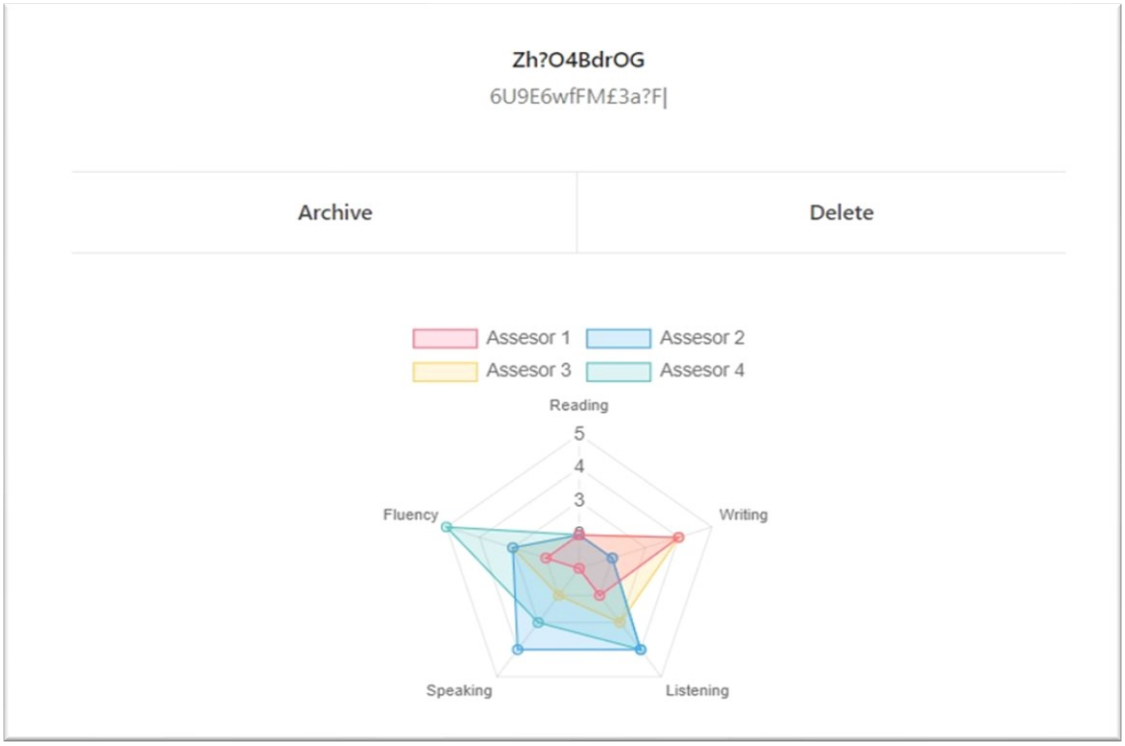



*FIGURE 14. Prototype 3. Radar chart made available to the person being assessed.*

# 5. DATA COLLECTION AND DATA ANALYSIS

This study presents data from various sources. The main source of findings is the focus group (FG) meeting that tested all the three Prototypes (see section 5.2. Recruitment procedure, demographics and background for the characteristics of the group). The author also obtained additional findings upon presenting Prototype 3 to his colleague (personal communication) and Prototype 2 to his MEE colleagues. These findings are called 'contextual', because they have provided extra context and paved the way for administering the FG. In all of the abovementioned instances, the participants received (a) a written invitation describing the nature of the research; and (b) a participant consent and data privacy form. Each of the participants was also asked whether they had any objections against being mentioned in the Acknowledgements section. Data collection from the FG is specifically addressed in section 5.3. below. The collection of contextual data is described in the current section.

***MEE workshop***

The MEE workshop (designed, hosted and organized by the author and co-hosted by one of the MEE peers) took place during the regular face-to-face (F2F) meeting at OAMK in February 2022. The primary aim was to sensitize the participants to the need for tools for giving structured feedback in student training. The secondary aim was to familiarize the MEE peers with typical issues that students of interpreting have during the early phase of their training. The personal aim was to fit-test the newly designed Prototype 2.

The participants received the invitations explaining the nature of the workshop and the variety of potential participant roles (Appendix 6). The participation was voluntary. There were nine active participants, one co-host and at least five observers. The design of the workshop (Appendix 7) was different from the FG, and in this respect their results are not directly comparable but complementary. Some of the author's peers from MEE had extensive background either in language teaching or interpreting, which is why the findings of this workshop are still relevant for this thesis. During the discussion the participants were asked, inter alia, to compare the efficacy of the two types of assessment. Namely, using no scaffolding vs using a tool with pre-defined categories. Three participants (Speakers) produced mock translations: essentially they repeated the text in English that was played to them. This activity was meant to mimic interpreting and reproduce its essential components, such as public speaking, memorizing bits of text, and producing a coherent narrative. Three participants (Assessors T-) evaluated the performance

based on the criteria of their choice. Three other participants (Assessors T+) evaluated the performance based on the tool that the author provided (Prototype 2). During the discussion at the end of the workshop the participants reflected on their experiences and the efficacy of the assessments. The workshop was recorded for further transcription in Otter.ai and coding in QDA Miner Lite. The coding frame with quotes is provided in Appendix 8.

**Personal communication**

During the period of data collection a number of colleagues of the author expressed interest in Prototype 3. This Prototype was seen as somewhat special for a few reasons. First, it was designed as a standalone web app for practice groups. Almost all of the colleagues were members of practice groups, and thus were interested in the new developments that facilitated training. Second, it was created at the author's request by a professional software developer. Because of this, it was perceived as a 'beta version' of a forthcoming product rather than an early 'alpha version', and naturally attracted more attention. However, due to imminent time constraints of the freelancers, it was impossible to conduct individual interviews and present the Prototype 3 to all of them. On the other hand, the author conducted one such interview with a freelancer colleague.

This was a 90-minute meeting in Zoom. Before that the written description of the interview and the consent form were provided. During the one-to-one interview, the colleague was offered the assessment exercise. First, a link to the prototype was sent to her. The author suggested that (for the sake of evaluating the user-friendliness of the prototype) the colleague tried out the interface with as little explanation from the author as possible. She found it very easy to handle and only needed minimal instructions. The author asked to assess the quality of simultaneous interpreting of a 5-minute speech from English into French using the arbitrary set of criteria. The author used the same speech from the EC (European Commission) meeting that was used for the FG. The simultaneous interpretation into French (one of the colleague's working languages) was also derived from the EC website. The colleague chose the criteria and their number at her discretion and entered them into the app. After that she requested to start both audio tracks and assessed the interpreter performance after listening. The assessment results were automatically saved in the app following the submission. After that the author asked three questions (one by one): "Could you talk me through the criteria that you chose?", "Now that you see the assessment result, what does it tell you?", and "Can you talk me through your experience with this tool: did you have any issues filling in the categories, etc.?" The answers were recorded in Zoom and later transcribed in Otter.ai. The results of this communication are described section 5.6. Personal communication results.

## 5.1. Qualitative data analysis procedure

*TABLE 1. Stages of data analysis, as implemented in this thesis*

| **Stages described in *Qualitative Content Analysis*** | **Implementation in this study** |
|---|---|
| 1. Deciding on a research question | The main research question and the corresponding sub-questions in this study: |
| | • RQ: What should be the characteristics of a software tool for developing core interpreter competences (common for consecutive and simultaneous interpretation)? |
| | o Sub-RQ1: What competencies should such tool support? |
| | o Sub-RQ2: Which characteristics should user interface of such tool have? |
| 2. Selecting suitable amount of material | Main data: |
| | - data derived from the FG meeting in April 2022. |
| | Additional data: |
| | - data derived from the workshop on structured feedback at OAMK in February 2022; |
| | - data derived from personal communication with a colleague interpreter. |
| 3. Building a coding frame | See Appendices 4 and 8 for details. |
| 4. Segmentation, i.e. fitting one unit of material into one (sub)category | Categories were checked for possible overlaps, similar categories were merged. |
| 5. Trial coding, i.e. applying the categories from the coding frame to the material (two rounds, 10 days apart if performed by one researcher) | All the data was subject to trial coding first followed by another coding session in 10 — 14 days. The categories from the trial stage were applied to the material and corrected accordingly. |
| 6. Evaluating coding frames and amending as needed | The coding frame was amended during the second round. |
| 7. Conducting the main phase of analysis, no more amendments to coding frame are allowed at this stage | The resulting coding frames were applied to the respective transcripts, so as to completely cover all the material. |
| 8. Presenting the coding frames and interpreting findings (coding frame itself is the result) | Coding frames and the interpretation of data are provided in the results section of this thesis. |

As Schreier (2014) puts it, the term 'qualitative content analysis' may be used to mean discourse analysis, conversation analysis, or thematic coding in different contexts. The author follows Schreier's version described in *Qualitative Content Analysis*, which is closest to thematic coding.

The procedure envisages building a provisional coding frame, testing it during the pilot phase in order to make the necessary amendments, and then going on with analyzing data using the definitive version of the coding frame. The systematic nature of the method is reflected in the fact that coding frames are unidimensional, mutually exclusive and, at least partly, data-driven (Schreier, 2014; see also Schreier, 2012). This study has implemented the eight stages of data analysis described by Schreier (2014), see table 1. Because coding frame is the cornerstone of thematic coding, the stage of Building a coding frame is subject to additional rules (Schreier, 2012):

- Selecting
    - selecting the parts of material that are relevant for the study;
    - creating a 'substantive coding frame' that is applicable to the selected material solely.
- Structuring and generating (i.e. deciding on categories and subcategories that exhaustively describe the selected data)
    - concept-driven, i.e. categories and subcategories based on prior knowledge, experience, theory or literature;
    - data-driven, i.e. categories and subcategories based on the material.
- Defining (i.e. explicitly defining the rules of assigning segments of data to categories)
- Revising and expanding
    - collapsing down to fewer categories if there are overlaps or too many similar sub-categories;
    - checking the coding frame for inconsistencies with the rest of the material.

In this thesis, the coding frame was based on the two research sub-questions, which were used as a constant comparative (Krueger, n.d.[29]; also *What is Constant Comparative Method*[30]).

[29] As Krueger puts it, "Analysis is sort of that: is that you're putting things into categories. In the literature we call this a *Constant comparative*. Actually this comes out of the writings of Barney Glaser and Absalom Strauss" (Krueger, n.d.).

[30] "Constant comparative method is a process developed by Glaser and Strauss and used in grounded theory, where you sort and organize excerpts of raw data into groups according to attributes, and organize those groups in a structured way to formulate a new theory" (Delve, n.d.).

All the codes had to be linked to either of the sub-questions:

(1) What competencies should the tool support?

(2) Which characteristics should user interface of such tool have?

Following the recommendation by Schreier (2014), the coding frame did not exceed three hierarchical levels. They were:

**For the MEE workshop:**

- Level 1: Concept-driven classes. The division produced by the author and based on practical experience, logic, common sense or theories. These are the most extensive clusters of responses. They include *Competencies* (*Self-reported* and *Observer-reported*) and *Experience giving feedback*.
- Level 2: Data-driven categories. Broad themes derived from the data that subsume the topics addressed by the participants. These themes reflect different aspects of *Concept-driven categories* that have been addressed during the workshop. In particular: *Self-reported competencies* include *Immediately reported competencies* and *Competencies reported on second thoughts*; *Observer-reported competencies* have the same subdivision; *Experience giving feedback* includes *Advantages* and *Disadvantages*.
- Level 3: Data-driven subcategories. Topics that the participants addressed. The list is extensive and is not provided here. It can be found in Appendix 8. The subcategories were mostly assigned using 'indicator' criteria, i.e. specific wordings contained in the quotes from the participants.

**For the FG and personal communication:**

- Hierarchy level 1: Classes. The division produced by the author and based on practical experience, logic, common sense or theories. These are the most extensive clusters of responses. They include: *Theoretical expectations* and *Practical observations*. *Theoretical expectations* include all the themes that the participants discussed before the actual assessment exercise, i.e. what they expected to be relevant in theory. *Practical observations* include all the themes raised after the assessment exercise, i.e. what has actually proven to be relevant to the participants (in their own opinion). In this case classes have mixed nature: they are both concept-driven and data-driven.
- Hierarchy level 2: Categories. Broad themes derived from the data that subsume the topics addressed by the participants. These themes reflect different aspects of *Classes* that have been addressed during the FG. In particular: *Theoretical expectations* include *Practical*

*skills*, *Personality traits* and *Aptitudes*; *Practical observations* include *Appeal to the assessor's mental model*, *Relative weight of criteria and the assessment context*, *The ease of handling the tool*, *The relevance of pre-filled prompts* and *Features to add*.

- Hierarchy level 3: Subcategories. Topics that the participants addressed. The list is extensive and is not provided here. It can be found in section 5.4. and in Appendix 4. The subcategories were mostly assigned using 'indicator' criteria, i.e. specific wordings contained in the quotes from the participants. Categories and subcategories also have mixed nature. Those categories and subcategories that were expected in theory are concept-driven. Those that were derived from the experience with the exercise are data-driven.

The rules of assigning classes and categories were developed as required by the procedure (Schreier, 2014). Subcategories were almost exclusively based on 'indicator' criteria (Schreier, 2012), which means that a particular quotation was assigned to a particular subcategory if it contained specific wordings.

### 5.2. Recruitment procedure, demographics and background for the focus group

This study relied on convenience sampling, meaning "targeting participants that are easy to find or access" (Harris et al., 2014, p. 76). One of the author's colleagues from Brazil (who also participated) invited six colleagues of her own to join the focus group (FG). As evident from table 2, all the participants are experienced professionals of interpreting. Some of them also have working background in translation and teaching. The participants have different educational backgrounds, and interestingly, only one person's original degree is in interpreting. The participant's real names have been altered to preserve their anonymity.

*TABLE 2. Focus Group participants, demographics and background*

| Name | Age | Gender | Educational background | Ethnicity | Current occupation | Native language | Working languages |
|---|---|---|---|---|---|---|---|
| Ashley | 35 + | female | Civil Engineering; certified translator | Brazilian | Interpreter/ Translator | Brazilian Portuguese | English, French, Spanish |
| Abigail | 35 + | female | History of Art | Brazilian | Interpreter | Brazilian Portuguese | English, French, Spanish |
| Audrey | 35 + | female | Law | Brazilian | Interpreter/ Translator | Brazilian Portuguese | English, French, Spanish |
| Annabelle | 35 + | female | Linguistics; phonology | Brazilian/ Swedish | Interpreter | Brazilian Portuguese | English, Spanish, Swedish |
| Alice | 35 + | female | Translation and Interpretation | Brazilian | Interpreter/ Translator/ Tutor | Brazilian Portuguese | English, Spanish |
| Arielle | 35 + | female | English language and Literature | Brazilian | Interpreter/ Translator/ Language Teacher | Brazilian Portuguese | English |
| Amanda | 35 + | female | Teaching English; Applied linguistics | Argentinian | Interpreter | Spanish | English, Brazilian Portuguese |

The invitations were sent to participants before the FG meeting. Eight persons took interest, however, one person could not attend. After that, a consent form and a pre-meeting questionnaire in MS Forms were sent to the participants. The FG meeting took place five days later. The goal of the questionnaire was to provide a better understanding of the discussion, to sensitize the participants to various theoretical and practical problems, and to collect the information about the participants' background knowledge and attitudes. The invitation form, the questionnaire and the consent form are provided in the Appendix section.

### 5.3. Data collection for the focus group

The FG meeting took place in Zoom and lasted 120 minutes. The audio of the meeting was recorded and further transcribed in Otter.ai. The remote (Zoom) format was chosen because of multiple locations of the participants (Spain, Finland, various cities in Brazil) and the ease of recording. During the meeting, the participants were asked to briefly introduce themselves and talk about their educational and working background. After that each participant was asked to talk the rest of the group through the most important skills and competencies that an interpreter should

have. The question was framed as "What would be the most important skills and/or competencies for an interpreter? Could you talk me through the main criteria that you think are the most relevant?" The first research question sought to identify the competencies that the practitioners of interpretation considered to be the most relevant. The question was first asked before the actual assessment exercise in order (a) to gain the understanding of the 'default' competencies that the interpreters keep in mind and, therefore, may want the new tool to promote; and (b) to get the benchmark for further comparisons.

After each of the participants had their say and made additions, the FG went on to the assessment part. A sample speech in English was offered which had to be assessed individually using Prototypes 1, 2 and 3. The assessment procedure (using all the Prototypes) required a recorded translation of a speech from one language into the other. Because different locations, institutions and individuals had been considered for a focus group since October 2021, the exact language pair was unknown. It was important to identify a publicly available source of speeches and their translations into various languages for a one-off session. Many repositories for interpreter training did not fit this purpose because provide monolingual speeches only (no translation). The website of the European Council / Council of the European Union (https://www.consilium.europa.eu/) seemed to be the optimal resource for the research purposes. However, simultaneous interpretation of meetings was available, which is why simultaneous mode was used.

The sharing of audios online has proven somewhat complicated. Normally, the assessment of simultaneous interpreting entails listening to two audio tracks at the same time (source language and target language). Obviously, the track volume has to be properly adjusted so that the assessor is comfortable listening. Apparently, sharing two tracks for 6 – 7 different people online would not allow individual adjustment of sound volume. For this reason, the author decided to offer 'staggered' or 'patched up' simultaneous: the audio tracks were split into several bits, approximately 3 – 5 sentences long. First came the English version and then the Spanish version. The participants were prompted in advance about this feature. In author's opinion, this modification would not complicate the assessment in a significant way.

However, some of the participants claimed that this design did not simplify listening or evaluation because it did not reproduce the typical assessment environment. They said that they would prefer to listen to both tracks at the same time in spite of the probable issues with the sound volume. The author shared both tracks at the same time via an audio editor, and this way the respondents assessed the simultaneous in the environment which many of them considered to be more realistic.

After the first assessment (Prototype 1) the participants were asked about the criteria that they were using. Then followed the second assessment of the same speech, now using Prototype 2. After this the participants were asked to compare their experiences with the two Prototypes, including, the experience with the user interface. The question about the assessment exercise was framed as "We've covered a list of skills and/or competencies during the first part of our meeting. Could you talk me through the criteria that you have used during the assessment part?" During these discussions, the topic of *Assessment situation should reproduce real-life situation* has emerged, which otherwise would probably have remained overlooked. The question to identify the user experiences with the Prototypes 1 – 3 was also asked after the actual assessment exercise and was framed as follows: "Could you talk me through your experiences with the user interface?"

Finally, Prototype 3 was offered. The participants commented on their experiences with the user interface. The list of categories/criteria that the participants used during this part of the exercise was anonymously saved in the app. This way, the comparisons of the categories became possible.

## 5.4. Focus group results

This section provides the results of qualitative content analysis. It shows the main classes (hierarchy level 1), categories (hierarchy level 2), and subcategories (hierarchy level 3). This section illustrates the descriptions with quotes and the frequencies of references. A detailed list of quotes with the rules of assigning categories and subcategories is provided in the Appendix section. This section describes the results of the FG meeting. The results of the MEE workshop and personal communication are shown in the following sections.

### 5.4.1. Theoretical expectations

Based on the process of coding described in section 5.1. Qualitative data analysis procedure, the following categories were identified: Practical skills, Personality traits, and Aptitudes. Practical skills encompass Language mastery, Thematic knowledge, and Professional demeanour. Personality traits include Being determined and Being mindful. Aptitudes include Willingness to learn and Being open to new developments.

#### 5.4.1.1. Practical skills

This category includes the practical know-hows that have been acquired through training. It incorporates Language mastery, Thematic knowledge and Professional demeanour. All the three aspects have received almost similar number of references. However, the approach to Professional demeanour among the FG participants was heterogeneous, which is why a part of the references was classified elsewhere (*Personality traits* and *Aptitudes* section).

**a. Language mastery**

Five out of seven interpreters identified language mastery as one of the top competencies. Two other interpreters did not seem to have any objections against what their colleagues have said, and chose to emphasize other competencies in the first place. Which probably may suggest that language mastery is considered self-evident among all the seven participants.

According to Ashley, an interpreter should have mastery both in their source and their target language. Ashley thinks that adequate language training may confer greater advantage over living in a native country. She uses her own example to support this point. Arielle and Abigail agreed with Ashley's point about the language mastery.

Audrey is of the same opinion. But on top of that she adds that mastering a language should be a matter of personal appeal rather than a duty, because this is something that should be done on a continuing basis. She says:

> *"Love the language that you are mastering. Mastering the language is important of course, you have to keep on learning forever and ever. So whatever your languages are, you have to work on them, read in them, listen to news in them, look for newspapers, news in them, read the articles."*

Alice gives a practical example of incorporating language enhancement into daily routines:

> *"Sometimes you're just watching the soap opera or a film [and you think to yourself]: 'how would I translate that in my language', or 'that expression is really interesting'."*

As evident from this subcategory, the FG largely recognizes the need for continuing development of language skills and thinks that it should become a second nature to interpreters.

### b. Thematic knowledge

All of the participants addressed thematic knowledge. Four persons pointed out their specific preferences, while three persons just enumerated their working topics. As shown in the *Recruitment procedure, demographics and background* section, most of the seven FG members majored in subjects other than interpreting/translation. However, they see it more as an advantage[31], because it gives more confidence working with particular topics. For example, Ashley says:

> "*I have a degree in Engineering, Civil Engineering, but I became an interpreter like 22 years ago, and that has been my profession to date. <…> Field knowledge is important. I mean, if you have a different degree, or any other experience working in the field, I think that helps. We all know that.*"

Audrey agrees and adds that she has special interest in environmental issues, which is where her degree in Law and an extension course in Environmental Management come into play and help her perform well. She also acknowledges that professional interpreters are "always plunged into different fields". Abigail has a degree in the History of Arts, which is one of her specializations:

> "*I graduated in History of Art and I hold a Master's degree in History of Art as well, which I've taught for 11 years <…> and I've been an interpreter since 1986. <…> [My] topics are Art, of course, Social and Human Sciences and Environment.*"

Interestingly, some of the FG participants who do not have any background in Law tend to refrain from this topic. Annabelle says:

> "*Most of my work has been in international cooperation and Business. And those are actually my favorite subjects. And also a lot of like all of us climate change, and now the pandemic and human rights and equity and gender issues, etc. I also do some medical, but it's not my favorite, and my 'hate subject' is Law. I really, really dislike working with Law*."

As Abigail put it, "I would work in other areas as well except for Law and nuts and bolts", the latter meaning Engineering. Meanwhile, Arielle, Alice and Amanda did not mention any specific preferences (or dislikes) that would be related to their professional/educational background, and just chose to enumerate the topics they most often work with.

[31] A different perspective on the effects of the available knowledge on problem solving is given in *Knowledge in Perception and Illusion* (Gregory, 1997).

This subcategory seems to suggest that, despite working with a wide range of topics in general, professional interpreters still have their favourite ones. These topics are often directly related to interpreters' educational background. Conversely, some topics in which interpreters have no background may become the least-liked (or even no-go) topics.

### c. Professional demeanour

All the participants specifically addressed the reliance on the professional patterns of behaviour rather than on the immediate instinct, maybe with an effort of will. However, this discussion had several aspects, which influenced the classification of responses. First, the 'technical' aspect emerged from the discussion. The participants addressed specific and relatively clear-cut techniques conducive to specific effects with listeners (such as keeping 'inner silence' that makes interpreter look confident; or mimicking the speaker's emotions that keeps the listeners of translation engaged, as described in the quotations below). Second, the 'intrinsic' aspect was discussed. According to some of the participants, high performance is a product of the pre-existing personal characteristics. And third, 'leveraging' aspect was brough up. The participants talked about exploiting various personal characteristics in order to develop useful reactions or habits that, in turn, help achieve good performance. The first group of answers is classified here, under the *Professional demeanour* subcategory. The second group is described in the *Personality traits* category, and the third in the *Aptitudes* category.

The *Professional demeanour* subcategory describes the responses of four participants. Two participants thought of *the reliance on professional behaviours* in terms of removing themselves out of the equation for the convenience of a particular client (i.e. producing 'zero' interference from interpreter[32]). Meanwhile, two other participants thought in terms of producing a predictable effect (i.e. factoring the interpreter in/producing 'positive' interference).

Abigail brought up the discussion by acknowledging the importance of putting ourselves in "other people's shoes" so as "to understand that their discourse is coming from somewhere else, not from my point of view".

Annabelle agreed and added that an interpreter should consciously shift their attention away from themselves and focus on their listeners instead. She describes this approach very much like a technique. In her own words:

[32] Cf. *Interpreting or Interfering* by Helen Tebble (Tebble, 2012)

> "*I try to take myself out of the picture and think that the person listening to me should be understanding the message that's been given as faithfully as possible to what's been given, rather than me trying to sit and figure out what is the best word or you know, not so self-centered and more to think about the person who's listening to me.*"

Annabelle recognizes that shifting the focus of attention can be difficult when the subject (topic) is complex and adds: "that's what I work on most of my time". This addition means that she is talking about a conscious technique in the first place. She also mentions the technique of preserving "the inner silence" once again when she talks about negative factors at work. While poor audio quality or failure to understand something is perfectly realistic in the interpreter's practice, this should not distract them too much. According to Annabelle, an interpreter should use the inner silence "as a compass in whatever you are doing".

Alice's approach is also very technical, but she tends to see the professional demeanour in terms of producing the same effect on the listener as the speech in the original does:

> "*If we can, we try to just mime in the style the person is using, so that we can really produce the same effect on the audience, so listening to the interpretation feels the same way to the audience as listening to the original audio, and the same reacting. So I think that this thing of putting yourself in the other person's shoes is very important.*"

Audrey picks up on this debate and also addresses the effects/impression:

> *"Train yourself to have nerves of steel or at least pretend that you do. Try not to crumble in the booth or with your colleague or especially if you're in a room full of people and you have a microphone in your hand. Try not to shake, that's important to give a good impression."*

In author's view, this subcategory reveals a very important variability of professionals' understanding of the professional approach to interpreting. While some interpreters see the professional behaviour as a product of learned attitudes and reactions (see the *Aptitudes* category below) that is very closely related to personality traits, others view it merely as a technique. This latter group of practitioners, in turn, also shows the heterogeneity of opinions. Although the wording "put ourselves in our clients' shoes" is common, it turns out to mean different things. For some interpreters, it means refraining from projecting their own impressions and feelings. For others, it nearly means creating a theatrical effect.

#### *5.4.1.2. Personality traits*

This category describes, following the definition of APA[33] (2007), the pre-existing internal characteristics that a person consistently shows in their behaviours, habits, attitudes, and feelings. These cannot be obtained through learning. The following subcategories are included: *Being determined* and *Being mindful*. This category includes the opinions of four out of seven FG participants.

**a. Being determined**

Solely the replies from two members of the FG were classified into this subcategory. However, it seems to highlight a very important difference, which is why it was created as a separate entity.

In order to achieve professional success, an interpreter needs, according to Ashley:

> *"To be a little bit bold, because sometimes we get there and we're not fully ready. So one must be bold. I'm talking about, you know, personality features."*

The latter point about the personality features does make a difference. Otherwise, we would have to classify this observation into *Being open to new developments* subsection of the *Aptitudes* section.

This point is further supported by Audrey who said that an interpreter should be capable of "a leap of faith" every time they start their assignment. Otherwise, they will be stuck in place and unable to proceed properly.

In author's view, this subsection reveals an important variability of professionals' opinions about 'being' or 'pretending to be'. Some participants think that pretending to be comfortable at work may be enough, others propose to actively derive this comfort through special techniques, while a few think that it should come naturally from the very start.

[33] American Psychological Association

**b. Being mindful**

Two other participants, Alice and Amanda, mentioned being deliberately aware of one's own feelings and mind so as to achieve the sense of calm. According to them, being endlessly harsh on oneself is counterproductive. In Alice's words:

> *"So I just want to say that we also have to be more indulgent on ourselves, we tend to judge ourselves too much. <…> So I think that we have to learn how to, you know, forgive ourselves for the mistakes, we are not perfect <…>, but many times we just tend to, you know, just sacrifice ourselves for many things. But when we just do one little mistake, we tend just to be sad for all the rest of the day."*

In her speech, Amanda explicitly referred to this point of Alice's and reframed it by saying that one has to be "humble" and "kind" to themselves.

#### *5.4.1.3. Aptitudes*

This category describes, following the definition of APA (2007), the potential to acquire particular skills or competencies through training. As was mentioned in the Professional demeanour subsection, a number of participants talked about the link between learned attitudes/ reactions and useful habits that eventually lead to high performance. These opinions are included in this category.

Aptitudes category incorporates the willingness to adapt and the openness to new developments. These subcategories characterize the ability to willfully rely on professional patterns of behaviour rather than on the immediate instincts. Unlike the Professional demeanour subcategory which describes focused techniques aimed at specific effects, Aptitudes describe more general attitudes that lead to broader effects.

**a. Willingness to learn and/or adapt**

Three participants of the FG emphasized that a professional interpreter has to be keen on learning and ready to readjust to the ever evolving working environment. Unlike the personality traits above, these characteristics can be obtained through learning (and participants seem to describe them as such). One of these characteristics would be flexibility. For example, Amanda said:

> "*I would just add flexibility. Because we really navigate in different settings. And that requires to be flexible. Many times, we don't have the ideal conditions. Yeah. And it's also different to behave in a setting where we are in a meeting with politicians, and you need to respect the protocol and be aware about these protocols. And when you are in the floor of a factory. So that requires a lot of flexibility.*"

In this case, being flexible means being ready to switch between different working environments and choose the most appropriate behaviour accordingly. This would not be a targeted technique (such as being focused on the listener or miming the style) conducive to a specific result (such as not distracting the listener or making the translation stylistically equivalent). Instead, Amanda seems to describe the leveraging of the pre-existing attitudes in order to promote better adaptability. This would seem to be one more type of adaptability on top of refraining from personal opinions.

For Audrey, the professional line of behaviour is about self-preparation. She says:

> *"And if you don't enjoy the research and learning about new stuff, it's terrible, and no, you can't do your work properly. So you have to enjoy the dive. Because for what from what I've read, and from what I learned in practice, I think for every hour in the booth you have at least two hours outside the booth studying, if not more. So if you don't enjoy the groundwork, don't do this job, it doesn't work."*

In other words, Audrey sees the adaptability to the working environment as a product of exploring different domains of knowledge. To paraphrase what she says, the natural curiosity can be consciously leveraged to deliver the proper level of performance. Natural curiosity cannot be learned, but exploiting it for the sake of interpreting can. She continues by adding that "you have to sorrow, no matter how tired you are". If you identify a piece of information that you are not familiar with, you have to explore it, even if you are tired. For Audrey,

> *"It's like school exams, and it will come up, so you have to learn it."*

In other words, delivering to standard is a matter of personal choice. An interpreter is free to either go the extra mile and research their subject thoroughly[34], or leave it where it is and risk to underperform. Because Audrey mentions the reliance on the existing habit to learn, her approach is different from the *Professional demeanour* subcategory above.

[34] Cf. "Desire without discipline dies… but desire with discipline deepens into passion… a discipline is an activity, within our power, that enables us to achieve indirectly what we cannot otherwise achieve by direct immediate effort. This is the difference between training and trying." (as cited in Fortis, 2020, p.6).

Meanwhile, Alice added the need for continuous learning. In her own words:

> *"And just as being a learner, not only fast learner, and not only someone who has a passion for learning, but I think that we are, we have to be an eternal learner."*

She also commented that the "passion for learning is a must". This would seem to add to Audrey's idea about exploring the different domains of knowledge.

This subcategory shows various avenues for interpreters to accommodate themselves to a wide range of work situations. These adaptations can range from the awareness of multiple contexts to proactive learning. All of the participants seem to emphasize how the existing attitudes, worldviews or habits can be employed in order to promote better professional performance.

### b. Being open to new developments

One more aspect of adaptability was addressed by four FG participants. The author classified it as *Being open to new developments*. This is a slightly different aspect, which is adaptability to the unexpected events. Arielle puts it this way:

> *"There's always some unknown. And we never know what to expect. We never know when we might come across a person whose line of thinking is not so clear. Or maybe we did not get aligned with the person or with the topic. So we always have to be ready for challenges. And this is part of our daily lives, and we are never ever ready."*

Annabelle raises the same point. She says: "think on your feet, be very, very flexible". Her comment picks up on Abigail's idea that one has to be a fast learner and fast thinker:

> *"A more technical aspect would be fast reasoning. Think fast, because what you are hearing might not be what you expect to hear."*

In author's prior personal communication, some colleagues called this trait "readiness to be imperfect" or "accepting the fact that you cannot be 100% perfect". This point was further developed by Alice:

> *"Sometimes we are just learning all new vocabulary as we go through all these slides of the speaker, right there on site and right there at the moment of the interpretation."*

That would be different from being a fast learner in the habitual circumstances. Ensuring 'basic minimum performance' despite the lack of customary resources was described by Christopher

Guichot de Fortis[35] from AIIC. He called this approach 'degraded mode' interpreting (Fortis, 2020). The opinions presented in this subcategory seem to reiterate this practical observation.

### 5.4.2. Practical observations

As was said at the beginning, the question about competencies was first asked before the actual assessment exercise in order to (a) reveal the 'default' competencies that the FG participants could think of on the spur of the moment; and (b) to get the benchmark for further comparisons.

When the same question was asked after the assessment exercise, this time, the following categories were identified: Appeal to the assessor's mental model, Relative weight of criteria and the assessment context, Personality traits, and Technical skills.

#### 5.4.2.1. Appeal to the assessor's mental model

This category describes instances when the FG participants mentioned differences or discrepancies between the 'real life' and the assessment situation, or between how they see the assessment situation and how the tool suggested to see it.

**a. The assessment situation should reproduce the real-life situation**

As mentioned in section 5.3. Data collection for the FG above, 'staggered' simultaneous was initially offered to the FG participants. This seemed to be a modification by necessity, although it differed from the traditional way of assessing performance in simultaneous interpretation. In author's opinion, this modification did not complicate the assessment in a significant way. However, three participants of the FG had a different view.

Annabelle said that, although she was prompted in advance, she felt like she did not assess the real situation. This made the assessment exercise difficult for her. She said that, for this reason, the exercise lacked purpose for her. She would prefer either 'genuine' simultaneous or 'genuine' consecutive, but not the 'artificial' version. As Annabelle put it:

[35] https://aiic.org/client/roster/clientRosterDetails.html?clientId=1094&clientRosterId=47&no_header=true&name=Christopher-GUICHOT%20DE%20FORTIS

*"You know, I'd rather either listen to both of them, at the same time, the way we do when we're in the booth or have a real consecutive interpretation, because I do not think that you can really take similar changes and transform it like this into consecutive."*

For Abigail, the modification has rendered the assessment exercise almost impossible, too. In her own words:

*"And I felt I wasn't able to evaluate it. Because actually, I was seeing it as a simultaneous interpretation in two parts. I wasn't able to see it as a consecutive. And it wasn't a real world situation in which we assessed the consecutive."*

Arielle addressed a different aspect of dissimilarity to the real life. Prototype 2 asked, among other, whether the speaker sounded relaxed and whether they were accurate. In Arielle's opinion, "we have to translate whatever comes our way", while the final result is not fully under our control. For this reason, the assessment of being relaxed or accurate is somewhat irrelevant.

This subsection demonstrates that a number of potential users may be extremely conscious of the categories that an assessment tool offers, or the similarity of the assessment situation to the real-life situation.

**b. The assessment tool should strike a chord with the assessor's mental model**

The criteria that the FG participants ended up using during the assessment exercise were different from the ones listed during the preliminary discussion. This section describes three opinions that shed light on the reasons behind this choice.

Ashley thought that there was "a mismatch" between the criteria that the FG has generated during the preliminary discussion, and the ones offered by Prototype 2. The assessment form asked about the speaker sounding "formal" or "engaging", which was not what the FG came up with at the beginning.

In Amanda's words,

*"I think that the evaluation was more regarding the language elements. Instead, we mentioned many things that have to do with personality and all that."*

This observation seems to suggest that the 'default' competencies that the interpreters consider in a theoretical discussion may differ from the 'assessment' competencies that they would be using in practice.

Audrey says that her answers were random because the assessment form did not seem to apply to what she was listening to. Furthermore, she said she would be more comfortable with summarizing the situation rather than providing a structured assessment using the particular set of the pre-defined criteria. She continues by emphasizing that the assessment criteria should have 'technical' nature:

> "*We just studied the ... what we heard on the tape. And as I said it ... it was said it became more technical. And for me, it was not what I was expecting for the form that you sent. I saw it in other eyes. For me, there was a discrepancy and all that.*"

This subsection shows the need for clear preliminary instructions for potential users. Otherwise, the discrepancies between the expected and the factual features of the tool may complicate user experience.

#### 5.4.2.2. Relative weight of criteria and the assessment context

This section reflects one of the most vividly debated elements of the FG discussion. When invited to give a free-form feedback about the interpreter performance, the participants actively engaged in considering the reasons why particular issues occurred. They also defined the order of priority of various criteria and their value. The *Relative weight of criteria* section describes the factors affecting the interpreter and the event. It includes three strong opinions that were expressed very persuasively and eloquently. Meanwhile, the *Assessment context* section describes six opinions about how the assessment procedure should be organized in order to be successful.

**a. Relative weight of criteria**

Audrey's observations regarded the intonation of the interpreter and the completeness of their delivery. In her opinion, the speaker was reading "at full cruise speed", while the interpreter tried to follow,

> "*And of course, he had to drop some sentences. And of course, some things were left out because there was no other way.*"

One can appreciate that, for Audrey, the assessment of completeness is only valid when the delivery of the original is adequate. As for the intonation, Audrey noted that

*"He tried to give some intonation where she didn't give much of any intonation. At some points, he tried to give some more. He presented his tone of voice, he tried to give it some intonation, but she didn't, she just read on."*

Although it is generally recognized that an interpreter should not be conspicuous by their presence (e.g. see the discussion about "removing oneself out of the equation"), Audrey seems to favour an exception in this case. Apparently, she approves of making the text sound more vivid in translation when the speaker is actually reading.

As was mentioned before, the interpreter on the recording made some omissions. In Alice's view,

*"you cannot really blame the interpreter for that. We don't know how … what kind of preparation he had, maybe this just wasn't something that came up on time, we know that many times our speeches come up like this, and you don't really have any, any idea of what the person is going to say."*

This point seems to add up to Audrey's observation: the validity of the assessment criteria should be defined by the quality of the input. In Alice's case, the input factors include the working environment, i.e. whether or not the interpreter received the materials on time.

Alice goes on by saying that in the assessment situation that she was offered, the interpreter should have been more precise, concise and focused. They should have disregarded all the rest parameters. In other words, she seems to be advocating the 'degraded mode' interpreting described by Guichot de Fortis (Fortis, 2012). Alice further supports her point by giving the example of her pedagogical work. When evaluating her students, she usually takes into account the type of speech, the style, and the quality of the input materials/prior preparation.

Ashley specifically addressed register. She said that the interpreter did not exactly follow the register of the original, but it would be appropriate in the case in question:

*"because I think the Latin language [meaning 'a romance language', such as Spanish] is you know, more fluid, a bit more relaxed."*

Apparently, the language pair becomes one more parameter to determine the relative weight of the assessment criteria.

This subsection reflects the general agreement among the FG participants about the fact that no criterion is absolute. On the contrary, the relevance of the assessment criteria is apparently subject to modification by a number of factors.

### b. Assessment context

This subcategory outlines the contextual factors that, according to the FG participants, affect the assessment. Six participants addressed this issue.

First, the FG participants recognized the context dependence of assessments in general. For example, Arielle said that [in the assessment using free forms] she had to amend the list of criteria several times. She was able to identify the final set of criteria arbitrarily and "on the go", while she was already listening to the audio. Amanda said that, although she completed the evaluation form [in the exercise using pre-defined categories], the assessment was "unfair", because she felt that some categories "didn't match or didn't apply at all".

Second, the FG participants highlighted the leading role of 'technical' approach in the assessment. In Ashley's opinion,

> *"when you are to actually assess an interpreter, as a trainer, you usually go to the more technical elements, more technical components."*

This was further evidenced by Audrey who also said that she "became more technical" when she heard the recording and had to assess it.

Third, although the principles of training in simultaneous interpreting (SI) and consecutive interpreting (CI) overlap[36], the participants still drew a distinction between the assessment of SI and CI. In Annabelle's opinion, the assessment of SI and CI "are two completely different things", and each requires "listening with a different ear". Furthermore, she added that CI assessment would require taking notes. Abigail fully agreed with Annabelle's opinion. Arielle also highlighted the importance of taking notes in CI assessment.

This subcategory testifies to the fact that the FG participants tend to draw a distinction between the 'general' and 'technical' skills/competencies. The 'general' and 'technical' skills/competencies have very few overlaps. Rather, the 'technical' skills/competencies may be viewed as the practical manifestation of the 'general' skills/competencies.

[36] See section 2.3. Development of Translation competences in the European region above.

### 5.4.3. What characteristics should the user interface have?

The coding procedure described in section 5.1. elicited three categories: The ease of handling the tool, The relevance of pre-filled prompts, and Features to add. Interestingly, the discussion about the user interface (UI) was lukewarm. This is in contrast with the MEE workshop that has seen extensive discussion of the UI. In author's opinion, it has to do with the nature of freelance work. A professional freelance interpreter works in many different locations, with different types of equipment and often with various colleagues. Although far from infinite, the selection of their topics is also vast. This tolerance to frequent changes and ad-hoc solutions probably explains the relatively weak interest in specific user interfaces. As previous sections have shown, professional interpreters seem to value much more the practicability of the tool's content and its adjustability to different contexts.

#### 5.4.3.1. The ease of handling the tool

This category reflects the participants' positive and negative attitudes toward the elements of the user interface. There are three subcategories: Freedom for users to determine criteria, Saving the results and Making the tool self-explanatory.

**a. Freedom for users to determine criteria**

The ability to choose the assessment criteria seems to play an important role. Although only two opinions have been voiced, it has to be noted that they were expressed very convincingly. Furthermore, other participants seemed to be in prefect agreement. Apparently, ther participants preferred not to expand on this topic because their opinion was likely identical to that of their colleagues.

> *"The Excel form? Yes. I was a bit lost about what criteria I should serve. But when I started listening, I changed the criteria, for instance, delivery. It was not among the criteria and so I added it, but well, it [Prototype 1] was easier than the second one [Prototype 2]."*

This way Abigail characterized the difference between the blank form and the pre-filled form. Being able to freely add criteria seems to be an advantage. Arielle is of the same opinion. She thinks that the blank form

*"gives us much more control, because we are going to choose what we are going to look at."*

On the other hand, she characterized the pre-filled form as "one size fits all". It may not necessarily fit a particular speech, which is a disadvantage.

**b. Saving the results**

Saving the results was mentioned by one participant, Abigail. Although the online Excel form did not require a specific saving procedure, this was not obvious. This has put Abigail "completely at a loss". Although this may seem an 'outlier' opinion, it obviously addresses an important piece of functionality and highlights a possible service design problem — all the more so because Prototype 3 actually has the 'Submit' button which needs to be pressed to save the results.

**c. Making the tool self-explanatory**

One participant, Amanda, pointed out that the prototypes and the corresponding workflow would benefit from additional explanations:

*"I think that for pedagogical reasons, it would have been useful just to go over the form, instead of just giving it to us and say, okay, this is it. We were really not prepared to what we were, what would we have to do? Yeah, then we realized that perhaps it didn't match quite well, the kind of setting we are given now."*

Accounted for the fact that Amanda is the only FG participant with formal education in English Language Teaching, this opinion has to be considered as a separate entity.

#### 5.4.3.2. The relevance of pre-filled prompts

This category has no further subdivisions. It describes the specific cases when the FG participants addressed the prompts offered by the tool.

Arielle reported having difficulty relating the pre-defined categories to the assessment situation. Although a preliminary discussion took place, it still was of modest help for the actual assessment. This probably testifies to the fact that the pre-filled prompts may not be efficient enough.

Annabelle addressed the same problem but in a greater detail. She enumerated the prompts one by one and commented if, in her opinion, they fitted the assessment. For example, she dismissed as irrelevant the following pre-filled criteria: “no need to listen twice”, “sounds just right, neither too academic, nor too informal”, “sounds relaxed and natural”, “the speakers are engaging”, “the speakers are keeping it short”. On the other hand, she highlighted as relevant the following criteria: “all the sentences are complete”, “the story sounds logical and coherent”, “names and numbers are credible and realistic”.

In author’s opinion, the findings of this section show the need for a very careful choice of wordings in prompts. Each person seems to gravitate toward their ‘personal language’, which is known as *idiolect*[37]. Apparently, when a prompt is in obvious contradiction with the user’s idiolect, the user may be likely to dismiss it as incorrect. Although this latter idea has received modest description in literature so far[38], the author is convinced that it should be more widely considered in designing software products/services.

If choice of wordings for prompts is made, ideally it should resemble the process of creating the CEFR descriptors — using the questionnaires based on multiple inputs: “So far data collection has been based chiefly on self-report <…> Nearly ten thousand respondents have completed questionnaires. <…> This is believed to be by far the biggest collection of data ever undertaken to validate a descriptive language proficiency scale.” (Council of Europe, 2021, p. 246).

#### 5.4.3.3. Features to add

This category has no further subdivisions. It describes the cases in which the FG participants talked about some features that, they felt, lacked.

Amanda has summarized various concerns about the to-add features. First of all, she addressed the need for clearer prompts. Second, she talked about the free-form feedback. The option of inserting a free text would be helpful. As well as the option of leaving out some points (non-mandatory fields). Furthermore, the stars feature in the Prototype 2 seemed questionable to her:

[37] Idiolect is “a term used in linguistics to refer to the linguistic system of an individual speaker — one’s personal dialect” (Crystal, 2008, p. 235).
[38] E.g. “John Potts asserts that the teacher’s own idiolect, the language choices they make which mark them out as individuals, plays an important part in the development of their students’ language ability. He makes a conscious effort to vary his own idiolect to give his students exposure to as large a ‘language bath’ as possible” (Gomm, 2013, passage 4).

*"I don't need to, for instance, mark any stars, if it is true stars, so it may not apply and perhaps include that as part of the script before starting the evaluation, something like that."*

In other words, the star symbol seems to convey a special meaning in the assessment. While a numerical value is viewed simply as a rank, a star probably characterizes positive impression. That is, number 5 may mean 'totally satisfied' vs number 1 'totally unsatisfied'. On the other hand, one star likely means 'I have some positive impression', but it will hardly ever stand for 'unsatisfied' at all. This finding seems to be in line with the previously made observation about idiolects. Apparently, some commonly used symbols may have a special meaning, at least for some of the users. Which is why the user interface of the proposed tool should at least allow users to choose the rating scales and the corresponding symbols.

### 5.5. MEE workshop results

This section presents the results of the MEE workshop that evaluated Prototype 2. As mentioned before, this data has contextual value. It adds up to the main (FG) findings and contrasts them. The data is presented in a three-component structure: main classes (hierarchy level 1), categories (hierarchy level 2), and subcategories (hierarchy level 3). The links between the MEE workshop data, the FG data and the Personal communication data are outlined in the Conclusions section. A detailed list of quotes with the rules of assigning categories and subcategories is provided in the Appendix section.

A workshop in structured feedback at OAMK in February 2022 explored the features that may be of interest for learners of languages, whereas the focus group meeting in April 2022 examined the applicability of the three prototypes for experienced practitioners of interpretation. Therefore, whereas EMCI students and recent graduates are the core target audience, experienced practitioners of interpretation and learners of foreign languages are extended-criteria audiences (see Figure 9 above).

As shown in Figures 10 and 11, Prototype 2 provides a set of nine questions framed in plain and almost colloquial language. The author's intention is to simplify the feedback process. Peers are expected to rate the rendition of the interpreter while listening. The ratings are assigned as stars (0 – 5) in response to the nine short questions about the interpreter performance.

Prototype 2 offers pre-defined categories. Users cannot make any changes to the categories or their number and can only grade performance. This Prototype is implemented as a poll in MS Forms. MS Power Automate is used to transfer the data from MS Forms to the Excel template. The template then builds radar charts.

On February 12, 2022 the author conducted a webinar on *Giving structured feedback* for MEE peers (Masters in Education Entrepreneurship). This was a group of post-degree learners with background in education and/or entrepreneurship. Many people were holding/ had held positions in administration or management. Several people were currently teaching. A few people were running their own business. The areas of expertise varied significantly from hospitality industry and language learning to project management and coding (programming). In general, the group was very active and willingly engaged in communication and practice-based exercises. A combined aim was to present some insights about structured feedback and, at the same time, to fit-test Prototype 2 with the extended-criteria target audience.

Based on the feedback received, the positive findings are:

(a) the participants thought that Prototype 2 promoted more open and comprehensive feedback;
(b) the participants believed that Prototype 2 promoted systematic assessment;
(c) the participants believed that Prototype 2 was user-friendly and easy to handle.

The negative findings are:

(a) the participants did not consider the numerical ratings to be optimal;
(b) the participants obviously lacked a free-text field for feedback;
(c) the participants did not find all the criteria to be self-explanatory.

In general, the webinar went as planned (see Appendix 10 for the plan), with many anticipated events happening (both anticipated difficulties and anticipated positive results). It also has brought a wealth of useful information summarized below.

**Anticipated events (and insights) of pedagogical value that have taken place:**

- User takeaways:
    - during the discussion, the participants have covered the most important criteria of feedback in interpretation, such as *Consistent messaging*, *Being familiar with a range of topics*, *Usage of backup strategies* (such as note-taking and reliance on visuals). These largely corroborate with the literature;

- some of the participants have praised the systematic nature of the tool. They've underlined the fast learning curve and user-friendliness. They've also underlined that the tool helps revise and compare user progress over time;

- several participants emphasized the need for meaningful numerical criteria and/or made a point that *a number is not always a proxy of performance*.

**Insights, downsides of the webinar and unexpected events:**

- User takeaways:
  - when asked about the competencies that the tool should support, several participants immediately addressed *Understanding targets of communication*. Understanding communicative targets and understanding the broader context of the situation is obviously an important skill, but it is often overlooked.

- Criticisms:
  - several participants have expressed a need for a free text field. Although assigning numbers to different criteria seems straightforward and simple, the demand for a 'personal qualitative comment' was present;

  - some of the participants have found simultaneous assessment difficult. Although rating several criteria on a scale of 0 to 5 during the performance (speech) seems straightforward, users still preferred to take some of their time before the assessment;

  - the fact that MS Forms were shown in a separate window was problematic for participants who used mobile phones. Although the broadcast in Zoom did not stop, they had to switch between different windows (or tabs) to make proper assessment.

This exercise has benefitted further prototyping. In particular, the author has made sure that his understanding of workflow with Prototype 2 was generally correct. The author has also enriched his understanding of important features and the scope of issues that need to be covered better in

communication. For example, participants who use mobile devices have to be informed about potential inconveniences switching between tabs.

In general, the participants have praised the tool for providing a systematic assessment framework. However, the ambiguous nature of some of the criteria and numerical assessments raises the need for preparatory instructions.

*TABLE 3. Summary of the MEE workshop findings*

| Research questions | Main classes | Categories | Subcategories |
|---|---|---|---|
| 1. What competencies should the tool support? | 1.1. Self-reported competencies | 1.1.1. Immediately reported | a. Understanding targets of communication |
| | | | b. Consistent messaging |
| | | 1.1.2. Reported on second thoughts | a. Being familiar with a range of new topics |
| | | | b. Usage of backup strategies (note taking, using meeting materials) |
| | 1.2. Observer-reported competencies | 1.2.1. Immediately reported | a. Being clear and concise |
| | | | b. Reproducing details in translation |
| | | 1.2.2. Reported on second thoughts | c. Overcoming 'technical' difficulties of the speech (names, numbers, adjectives) |
| 2. What is the user experience with the tool? | 2.1. Advantages | 2.1.1. Features which are helpful in general | a. Fast learning curve |
| | | | b. Systematic assessment is possible |
| | 2.2. Possible improvements | 2.2.1. Repeating in a different environment | a. To repeat with more assessors |
| | | | b. To repeat the assessments after learners have already familiarized themselves with the tool |
| | | | c. To compare oral feedback and tool-based feedback from the same users |
| | 2.3. Disadvantages | 2.3.1. A minus but not a barrier | a. Lack of free text field |
| | | | b. Need for several rounds of pilot tests |
| | | | c. Assessment implemented simultaneously |
| | | 2.3.2. Definitely a barrier | a. Criteria not always clear |
| | | | b. Assessments are numerical |
| | | | c. Forms and webinar shown in different windows (need to switch) |

### 5.6. Personal communication results

This section briefly outlines the results of presenting Prototype 3 to a colleague interpreter. The findings are listed as a series of bullet points. The aim of this section is to provide additional context for the discussion. The links between the MEE workshop data, the FG data and Personal communication data are outlined in the Conclusions section.

The colleague, whose responses are presented here, is a freelance interpreter/translator with more than 10 years' experience, including, at various international organizations, such as specialized agencies of the United Nations. She has strong educational background in English philology as well as Conference Interpreting and Translation. The results of the interview are provided below. The colleague will be referred to as Ivonne, which is not her real name.

***Could you talk me through the criteria that you chose?***

- Before the assessment Ivonne listed the following groups of criteria:
    - Group 1: Delivery; Pace; Volume/tone/intonation; Breathing/pauses;
    - Group 2: Target language quality; Grammar; Vocabulary/Syntax;
    - Group 3: Accuracy; Main message; Details; Summarization/Recasting.
- After listening to the audio, she said that Accuracy, Delivery and Quality of the target language would be enough.
- According to the colleague, every meeting may warrant a set of individual and very specific criteria, which is why user input is obviously better than any pre-defined lists.

***Now that you see the assessment result what does it tell you?***

- Ivonne said that she would prefer a more detailed assessment scale: namely, one with 0.5-point increments (rather than with 1-point increments).
- She liked the sliders better than any other alternative. When asked about stars, Ivonne replied that "stars give me a consumer assessment perspective. I just get the associations with reviewing a physical product, which is not appropriate in the given context".
- As for the radar chart, the expert said that it would probably be more informative with many inputs and multiple overlaps. She said that with fewer assessors she would prefer

a conventional bar chart, whereas with many assessors she would prefer the radar chart. In Ivonne's opinion, the usability of the radar chart that Prototype 3 offers is limited to summarizing feedback from multiple assessors and/or several sessions.

***Can you talk me through your experience with this tool: did you have any issues filling in the categories, etc.?***

- Ivonne said that she was able to do the assessment only after she had listened to the whole recording. The reason was that she wanted to see the audio track as well, but did not have two screens.
- She did not like the authentication implemented in the app. The combination of random symbols looked unnatural to her, and was a bit of a nuisance.
- Ivonne said that the division into categories and subcategories of assessment got her confused during the exercise. As she put it, "I couldn't see a clear distinction between a set and a subset".
- The expert also noted that she intuitively felt like saving the results, but could not find the 'Save' button.
- According to Ivonne, the position of the knob on the assessment slider (graded from 0 to 5) "was somewhat biased". As she put it, the knob "is positioned towards a higher grade" [it is set at '3' by default, which is closer to '5' than to '0'].

These observations add up to the input from the FG. Although the protocol of this conversation was somewhat different from that of the FG, it still provides relevant information. Furthermore, the professional and educational profiles of Ivonne were almost identical to those of the FG participants, which further justifies at least limited comparisons of the findings.

# 6. Conclusions and reflection

This study aimed to answer the following question: 'What should be the characteristics of a software tool for developing core interpreter competencies (common for consecutive and simultaneous interpretation)?' Two sub-questions were: (a) 'What competencies should such tool support?'; and (b) 'Which characteristics should the user interface of such tool have?' This section summarizes the answers to these questions based on inputs from three sources: focus group meeting (main source; hereafter FG), MEE workshop (ancillary source; hereafter MEE), and personal communication with a colleague interpreter (ancillary source; hereafter PC). It also draws parallels between these findings and the theoretical background described in Chapter 2.

As was mentioned in the Presentation of portfolio works (Chapter 4), the author presented Prototypes 1 – 3 to the FG. He also presented Prototype 2 to his MEE peers and Prototype 3 to his colleague interpreter. This has provided extra context to contrast the findings of the focus group and has paved the way for the prototyping procedure. The outcomes of these discussions are arranged into two groups (corresponding to the research sub-questions) with further subdivisions based on the commonalities identified.

## 6.1. Characteristics of the tool (competencies)

This section summarizes the research findings that answer the first research sub-question. As evident from the following subsections, a tool for developing core interpreter competencies should be:

- consistent with the demand-driven approach;
- consistent with the aim of trainees' entering into profession;
- consistent with the CEFR-like approach in Translation;
- consistent with the aim of critical thinking about the Translation process.

Although the inventory of specific competencies and their relative weights vary, a generally recognized list of invariant competencies may include:

- mastery of the target language;
- accuracy of messaging; and
- coherence of presentation.

The overview of these observations is provided in detail below.

### 6.1.1. Consistency with the demand-driven approach and context-dependence

One of the themes that has proven common for FG, MEE and PC is the consistency of the tool with the demand-driven nature of Translation. The MEE voiced the importance of understanding targets of communication. PC pointed out that the assessment of any meeting warrants a set of specific and individual criteria. Meanwhile, the FG has recognized the dependence of assessments from the context in general. The FG has also reiterated the concept of 'degraded mode interpreting' and made a case for the relative weights of criteria based on circumstances. These observations would be in line with the trend for demand-driven approach in interpreter training, such as the usage of event-specific materials, focus on the effective management of event-specific information, or critical thinking about the problems pertaining to the profession. The findings from the FG and the PC have shown that professional interpreters seem to object against any sort of scaffolding in general, whereas the MEE did not dismiss the scaffolding altogether. Therefore, if the tool were to offer some type of scaffolding (drop-down lists, pre-defined criteria, pre-arranged blank fields, etc.), this framework should be very flexible, customizable and adaptable to the user needs.

### 6.1.2. Orientation to entry into profession

One of the FG and MEE findings that corroborates with the theoretical framework is the professional orientation of the tool. As described in Chapter 2, Daniel Gile (2009) raised the need for ad-hoc acquisition of knowledge. The Comillas curriculum highlights the importance of familiarity with international issues of the day. Meanwhile, MEE addressed the importance of being familiar with a range of topics, and the FG made a case for the advantages of having other degrees on top of the degree in interpreting/translation. Therefore, if the tool were to offer any type of scaffolding, it would have to take into account different knowledge domains and, probably, variable structure of knowledge.

### 6.1.3. Consistency with the CEFR-like approach and proficiency levels

The trend for CEFR-like approach and proficiency levels in Translation were described in Section 2.1. Principles of the professional practice. The work toward the competence levels framework in Translation is underway, and it strengthens the role of some invariant competencies. The inventory includes, in particular, the mastery of the target language, accuracy of messaging

and coherence of presentation. The criteria that MEE and FG suggested would seem to corroborate with these. In particular, MEE highlighted such competencies as being clear and concise, and reproducing details in Translation, including names, numbers and adjectives. Meanwhile, the FG shared an extensive list of criteria, of which the most common ones were accuracy and coherence. Furthermore, the FG was unanimous in recognizing the role of mastery both in the source and in the target language. Therefore, the conceptual framework of the tool needs regular synchronization with the competence levels framework. The FG findings have shown that professional interpreters tend to draw a distinction between the 'general' and the 'technical' skills/competencies. For this audience, the skills/competencies that the Prototypes incorporate would be relevant solely in the context of formal assessment. The 'general' and the 'technical' skills/competences have very few overlaps. Rather, the 'technical' skills/competences may be viewed as the practical manifestation of the 'general' skills/competences. Therefore, the expert system that the author is designing should propose at least two modes of operation: training mode vs assessment mode.

### 6.1.4. Orientation to the process of Translation

As described in subsection 5.4.2.2. Relative weight of criteria and the assessment context, the assessment of interpreter performance in the FG was largely centered around the reasons *why* the interpreter performed in a particular way rather than *what rating* should be assigned to their performance. This shift of attention from '*what'* to '*why'* was described in subsection 2.2.3. Process oriented approach. Such a shift is seen as conferring greater efficiency to interpreter training.

Although most of the FG members originally majored in subjects other than interpreting/translation, the attention to '*why'* was still highly relevant to the FG. In author's opinion, this mindset could have been shaped by empirical observations, professional communication, or continuing professional development. This would probably mean that the attention to '*why*' is nearly ubiquitous in interpreter/translator training, and should be taken into account for the development of the tool. The audience of professionals would probably benefit from the tool if a 'Field Guide' was added that explains the reasons why some observed phenomena are taking place in interpretation.

### 6.2. Characteristics of the tool (user interface)

This section summarizes the research findings that answer the second research sub-question. The review of the tools to support competence development in interpreting/translation (Chapter 3) has shown that important characteristics include:

- account for patterns of delivery;
- account for content;
- account for negative feedback.

Furthermore, the user interface of such tools should allow: asynchronous assessment, custom feedback and time-coding/linking with the source.

One of the most noteworthy observations is that the user feedback from MEE peers and the PC/FG respondents (professional interpreters) was quite different, although with some overlaps, both in terms of its structure and its content (see Appendices 4 and 8). The two groups addressed different points and highlighted different concerns. Interestingly, same questions generated variable interest among the two groups. For example, both were asked about the user interface. As evident from Appendix 8, the MEE group has covered this topic extensively, whereas the FG addressed this matter very briefly and preferred to expand on the context-dependent nature of assessment and the factors that determine the relative importance of different criteria.

Some of the author's MEE peers can be viewed as an extended-criteria target audience: a few of them have first-hand experience with English language teaching, which is deeply rooted in CEFR (Common European Framework of Reference for Languages). As was mentioned in Chapter 2, recent years have seen a considerable momentum for introducing a framework for teaching interpretation which would be similar to CEFR. Therefore, the structure and the principles of the assessment that the Prototypes provide may be valid for language teaching as well.

To sum up, these observations have brought the author to conclude that the role of the user interface will be different for the two target audiences. Language teachers would want free text fields, variable assessment scales (not always numerical), a separate window to run the assessment software, etc. While similar points were stressed by interpreters, they placed a much greater value on the functionality that would allow them to assign different weights to the assessment criteria and to create different assessment profiles for different contexts.

### 6.2.1. Allow asynchronous assessment

The results from the FG, MEE and PC have shown that grading performance while listening may be undesirable for a number of reasons. One reason would be the willingness to concentrate on the interpreter's rendition and get general impression first, rather than to start off with structured assessment straight away. Another reason would be the insufficient clarity of criteria for the assessors (e.g. in the context of peer-assessment) or their unwillingness to rate the performance against specific criteria. One more reason would be remote learning that often relies on asynchronous modalities of assessment. Some of the commercially available tools for interpreters, such as GoReact, incorporate the asynchronous functionality, which may additionally testify to its relevance.

### 6.2.2. Allow custom forms of feedback

The need for free text fields has been raised by FG, MEE and PC alike. This probably has to do with the fact that the assessment procedure is context-dependent and should take into account at least the type of Translation, the degree of similarity of the training environment to the real-life situation, and the applicability of the assessment criteria. Furthermore, a number of external factors can modify the relative importance of various assessment criteria, which users may want to reflect in free-feedback forms.

Therefore, the proposed tool should offer (a) free text fields; and (b) different presets or profiles for various types of texts and assessment situations, including, the modifiable lists of criteria, and the option to assign relative weights to the criteria.

### 6.2.3. Allow time-coding / linking with the source

One of the difficulties reported by FG, MEE and PC was the need to open the source audio/video and the assessment app in different windows and, consequently, to switch between the two windows. Although it was the author's responsibility to adjust all the audio/video controls in the context of this study, participants still wanted to have these controls within their reach. This gives an important insight into the design of the user interface. Furthermore, time-coding / linking with the source is promoted as one of the competitive advantages of GoReact, which may additionally testify to the high relevance of this option.

The author's overall impression is that the three Prototypes have been relevant to the target audience of interpreters. The more flexible and user friendly Prototypes 2 and 3 (MS Forms based and TailwindUI based, respectively) have apparently performed better than the plain Excel table. Although easy to use and visual, these prototypes are only the first step in a series of further iterations toward a truly individualized expert system. The functionality of such system should be enriched with more features. In order to identify the set of criteria and their relative weights, more focus groups will be required. The preparation of these focus groups should take into account the concept of *strategic approaches to learning* offered by the ASSIST questionnaire (Tyler, 2007; Skarlatos, 2021). Furthermore, the concept of *epistemic profiles* (e.g. Lonka, 2021) has to be explored in order to account for the context-dependent nature of the assessment criteria in interpreting.

### 6.3. Contributions of this study

This study presents the Prototypes of a lo-tech, visual and customizable tool. It can either be implemented as a standalone web-app or be based on Microsoft suite. This approach offers a budget ad-hoc solution for collaborative learning.

The findings of the FG have shown that the tool can be deployed and used relatively easily. Furthermore, important differences across the target audiences have been found, which gives insights into the market segmentation.

### 6.4. Limitations

One of the limitations of this study would be its context-dependent nature. It is a qualitative study with a limited number of participants and non-random sampling. Therefore, its results cannot be generalized to a broader audience representing wider age and gender groups. Although the coding was conducted in two rounds at least 10 days apart, as recommended for studies performed by one researcher (Schreier, 2012), peer-validation or triangulation of the results would be beneficial. Furthermore, the study explored the core competencies that are common for consecutive and simultaneous interpretation. However, a more specific analysis of needs in consecutive and simultaneous interpretation would be useful.

# 7. References


A Visual Summary: 32 Learning Theories Every Teacher Should Know. (n.d.). Retrieved from TeachThought: https://www.teachthought.com/learning/a-visual-summary-the-most-important-learning-theories/?

Afonin, S. (2011). Shaping Translator's Competences for the United Nations Russian Translation Service. In Shipinov, I. Training the Next Generation of Professionals: The St. Petersburg School of Conference Interpreting and Translation (стр. 9 – 28). St. Petersburg: Herzen University Press.

AIIC. (2004). AGREEMENT ON WORKING CONDITIONS AND THE PECUNARY REGIME FOR CONFERENCE INTERPRETING AGENTS1 (ACIs) RECRUITED BY THE INSTITUTIONS OF THE EUROPEAN UNION. Retrieved from https://aiic.org/document/4905/AIIC-EU-Convention-ENG.pdf

AIIC. (2021, 10 12). Best practice guidelines for conference interpreting training programmes. Retrieved from AIIC: https://aiic.org/site/world/conference/becomingci/bestpractice

Alekseeva, I. S. (2011). Post-graduate translator/interpreter training programme at the St.Petersburg School of Conference Interpreting and Translation. In Shipinov, I. Tranining the Next Generation of Professionals: The St.Petersburg School of Conference Interpreting and Translation (стр. 34 – 37). St. Petersburg: Herzen University Press.

Alekseeva, I., & Antonova, A. (2011). Experience of the St.Petersburg School of Conference Interpreting and Translation in Coordinating the Curriculum with Pedagogical Assistance Provided by the UN under the MOU. In Shipinov, I., Traning the Next Generation of Professionals: the St.Petersburg School of Conference Interpreting and Translation (стр. 29 – 34). St.Petersburg: Herzen University Press.

Baker M. (2008). Routledge Encyclopedia of Translation Studies. London: Taylor & Francis Group.

Beeby, A. (2009). Results of the validation of the PACTE translation competence model: Acceptability and decision making. Across Languages and Cultures , 207-230.

Boulton, A. (2017). Corpora in language teaching and learning. Cambridge University Press. Available at: https://www.cambridge.org/core/journals/language-teaching/article/abs/corpora-in-language-teaching-and-learning/FFB14364C16ABDD0DD78796D197BBDB5.

Cambridge Assessment English. (2019). Cambridge English Teaching Framework Components. Retrieved from Cambridge Assessment English: https://www.cambridgeenglish.org/Images/172991-categories-and-components-cambridge-english-teaching-framework.pdf

Council of Europe. (2021, 10 15). Common European Framework of Reference for Languages: Learning, teaching, assessment. Retrieved from Council of Europe: https://rm.coe.int/1680459f97

Council of Europe. (2021, 10 15). Common European Framework of Reference for Languages: learning, teaching, assessment. Companion volume. Retrieved from Council of Europe: https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4

Crystal, D. (2008). A dictionary of linguistics and phonetics. John Wiley & Sons, Incorporated.

Delve. (n.d.). What is Constant Comparative Method? Retrieved from Delve: https://delvetool.com/blog/ccm

Demidov, M. (2020). Teaching Interpretation: Progress Assessment Toolkit (English version). Retrieved from Zenodo: https://zenodo.org/record/3755784#.Yi42nHpBxM2

Demidov, M. (2020). There is more functional language around than you may think. IATEFL VOICES, 19.

Demidov, M. A. (2019). The Construction And Activation Of Language Users' Background Schemata In Medical Domain Discourse. The Journal Of Teaching English For Specific And Academic Purposes, 557 – 572.

Deysel, E. (2018). Experimenting with computer-assisted interpreter training tools for the development of self-assessment skills: National Parliament of RSA. In C. F. (Ed.), Interpreting and technology. Translation and Multilingual Natural Language Processing. Berlin: Language Science Press.

DG Interpretation. (2021, 11 5). Planning and management documents Strategic Plan 2020-2024. Retrieved from An official website of the European Union: https://ec.europa.eu/info/system/files/scic_sp_2020-2024_en.pdf

eLearning Africa (ELA). (2022, March 30). Interview with Keynote Speaker Mark West. Retrieved from ELA Newsportal: https://ela-newsportal.com/interview-with-keynote-speaker-mark-west/

EMCI. (2021, 10 10). Welcome to the European Masters in Conference Interpreting! Retrieved from EMCI: https://www.emcinterpreting.org/

European Commission. (2017). Retrieved from Digital Competence Framework for Educators (DigCompEdu): https://joint-research-centre.ec.europa.eu/digcompedu_en

European Commission. (n.d.). Books and E-books. A bibliography of books and e-books on conference interpreting. Retrieved from An official website of the European Union: https://ec.europa.eu/education/knowledge-centre-interpretation/conference-interpreting/research-and-background-reading/books-and-e-books_en

European Commission. (n.d.). Consecutive Interpreting. What is consecutive interpreting? What does it entail? Retrieved from An official website of the European Union: https://ec.europa.eu/education/knowledge-centre-interpretation/conference-interpreting/consecutive-interpreting_en

European Masters in Conference Interpreting. (n.d.). EMCI Core Curriculum. Retrieved from EMCI: https://www.emcinterpreting.org/emci/core-curriculum

Fantinuoli, C. (2009). InterpretBank: Ein Tool zum Wissens-und Terminologiemanagement für Simultandolmetscher. *Tagungsband der internationalen Fachkonferenz des Bundesverbandes der Dolmetscher und Übersetzer eV (BDÜ)*, 411-417

Fantinuoli, C. (2016). Proceedings of the 38th Conference Translating and the Computer. InterpretBank. Redefining computer-assisted interpreting tools, (pp. 42 – 52). London.

Fantinuoli, C. (2017). Computer-assisted preparation in conference interpreting. Translation & Interpreting, 24 – 37.

Felder R., B. R. (2005). Understanding Student Differences. Journal of Engineering Education, 57-72.

Felder, R. (2021, 10 26). Index of Learning Styles Questionnaire. Retrieved from ResearchGate: https://www.researchgate.net/publication/228403640_Index_of_Learning_Styles_Questionnaire

Felder, R. M. (1988). Learning and Teaching Styles in Engineering Education. Journal of Engineering Education , 674-681.

Finkelstein, S., Fishbach, A. (2012). Tell Me What I Did Wrong: Experts Seek and Respond to Negative Feedback. Journal of Consumer Research, 22 – 38.

Flowerdew, J. (1996). Concordancing in Language Learning. In M. C. (Ed.), The Power of CALL (pp. 97 – 113). Houston: Athelstan.

Fonseca, D. (2022). Online interpreting practice groups: your buddy can be anywhere! Interview with Daniele Fonseca. Retrieved from https://www.youtube.com/watch?v=aKGZr4E-HyQ

Fortis, C. G. (19 June 2020 r.). Advice for conference interpreters who are struggling in the booth: how to switch into 'degraded mode' to ensure 'basic minimum' performance. Retrieved from Cambridge Conference Interpretation Course: http://www.cciconline.net/documents/texts/Degraded%20mode%20(EN)%20-%20final.pdf

Gambier Y. (2007). Doubts and Directions in Translation Studies : Selected contributions from the EST Congress, Lisbon 2004. Amsterdam: John Benjamins Publishing Company.

Gaskell, D. (2004). Can learners use concordance feedback for writing errors? System, 301 – 319.

Gile, D. (2009). Basic concepts and models for interpreter and translator training. Amsterdam: Benjamins Publishing Company.

Goldsmith, J. (2022). Build an awesome glossary in no time flat. Retrieved from 2022 Innovation in Interpreting Summit: https://iiisummit.squarespace.com/s/2022-Innovation-in-Interpreting-Josh-Goldsmith-Glossaries-Worksheet.pdf

Gomm, H. (2013). Issue 83: idiolects; teacher personality and more. Retrieved from Modern English Teacher: https://www.modernenglishteacher.com/issue-83-idiolects-teacher-personality-and-more

Gregory, R. L. (1997). Knowledge in perception and illusion. Philos Trans R Soc Lond B Biol Sci, 1121–1127.

Harris, M., Taylor, G., Jackson, D. (2014). Clinical Evidence Made Easy. Scion Publishing Limited.

Horváth, I. (2016). Information and Communication Technologies in Interpreting and Machine Interpretation. The Modern Translator and Interpreter, 183–189.

Hurtado Albir, A. (2017). Researching Translation Competence by PACTE Group. Amsterdam: John Benjamin's Publishing Company.

Hurtado Albir, A. (2018). Competence Levels in Translation: Working Towards a European Framework. Interpreter and Translator Trainer, 111 – 131.

Instituto Cervantes. (2018). DELE C1 CUADRO DE ESPECIFICACIONES . Retrieved from Instituto Cervantes. Exámenes.: https://examenes.cervantes.es/sites/default/files/DELE-C1_Especificaciones-2018.pdf

Jarvis S., Daller, M. (2013). Vocabulary Knowledge : Human Ratings and Automated Measures. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Krueger, R. (n.d.). Focus Group Session 4 – Analysis. Retrieved from https://www.youtube.com/watch?v=5sMe_iQfd0s&t=966s

Learning Theories and Models - Summaries & Guides. (n.d.). Retrieved from https://www.learning-theories.com/

Littig, B., & Pöchhacker, F. (2014). Socio-Translational Collaboration in Qualitative Inquiry: The Case of Expert Interviews. Qualitative Inquiry, 20(9), 1085–1095. https://doi.org/10.1177/1077800414543696

Lederer, M. (2003). Translation : The Interpretive Model. London and New York: Taylor & Francis Group.

Lonka, K. (2008). MED NORD - a tool for measuring medical students' well-being and study orientations. Medical Teacher, 72-79.

Lonka, K. (2012). How to make a lecture course an engaging learning experience? Studies for the Learning Society, 63-74.

Lonka, K. (2018). Phenomenal Learning from Finland. Helsinki: Edita.

Lonka, K. (2021). University students' epistemic profiles, conceptions of learning, and academic performance. Higher Education , 775–793.

McEnery, T., Xiao, R. (2011). What Corpora Can Offer in Language Teaching and Learning. In E. H. (Ed.), Handbook of Reasearch in Second Language Teaching and Learning (pp. 364 – 375). Routledge.

Navigli, R. (2012). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. Artificial Intelligence, 217 – 250.

Navigli, R. (2021). Ten Years of BabelNet: A Survey. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21), 4559 – 4565.

Newell, S. (2002). Behaviour Management in the Classroom. A transactional analysis approach. London: David Fulton Publishers.

Nieto, M. (Ed.). (2014). Corpus-based translation and interpreting studies : From description to application : estudios traductológicos basados en corpus: de la descripción a la aplicación. Frank & Timme.

O'Donnell, A. (1987). Cognitive, Social/Affective, and Metacognitive Outcomes of Scripted Cooperative Learning. Journal of EducaUonal Psychology, 431-437.

Orlando, M. (2016). Interpreting training and digital pen technology. Use of the Smartpen opens up new possibilities in interpreting training and practice. Retrieved from https://aiic.org/document/885/AIICWebzine_2013_Issue62_8_ORLANDO_Interpreting_training_and_digital_pen_technology_EN.pdf

PACTE Group. (2018). Competence levels in translation: working towards a European framework. The Interpreter and Translator Trainer, 111 – 131.

Papageorgiou, S. T. (2015). The Association Between TOEFL iBT®Test Scores and the Common European Framework of Reference (CEFR) Levels. Princeton, New Jersey: Educational Testing Service.

Pöchhacker, F. (2001). Quality Assessment in Conference and Community Interpreting. Meta, 410–425.

Pöchhacker, F. (2014). Assessing aptitude for interpreting: The SynCloze test. In M. L. Franz Pöchhacker, Aptitude for Interpreting (pp. 147 – 161). Amsterdam / Philadelphia: John Benjamins Publishing Company .

Poger, J. (2018). The Business of Interpreting: FAQ 9 – What is the customer decision journey? . Retrieved from https://aiic.org/document/1055/AIICWebzine_2018_Issue73_11_POGER-GUICHOT_DE_FORTIS_The_Business_of_Interpreting_FAQ9_What_is_the_customer_decision_journey_EN.pdf%20right%20click%20and%20Copy%20Link%20Location

Poger, J. (2018). The Business of Interpreting: FAQ 9 – What is the customer decision journey? . Retrieved from https://aiic.org/document/1055/AIICWebzine_2018_Issue73_11_POGER-GUICHOT_DE_FORTIS_The_Business_of_Interpreting_FAQ9_What_is_the_customer_decision_journey_EN.pdf%20right%20click%20and%20Copy%20Link%20Location

Rütten, A. (2017). Interpreting and the Computer – finally a happy couple? Retrieved from Dolmetscher wissen alles: https://blog.sprachmanagement.net/interpreting-and-the-computer-39/

Schreier, M. (2012). Qualitative Content Analysis in Practice. SAGE.

Schreier, M. (2014). Qualitative Content Analysis. In Flick, U. The SAGE Handbook of qualitative data analysis. SAGE.

Shutova, E. (2011). How to Hit the Ground Running: Translation and Interpretation as a Profession. In Shipinov, I., Training the Next Generation of Professionals: The St.Petersburg School of Conference Interpreting and Translation (pp. 47 – 48). St. Petersburg: Herzen University Press.

Skarlatos G. Dedos, D. F. (2021). Dataset and validation of the approaches to study skills inventory for students. Scientific Data volume.

Smith, A., Harris, S. (2014). Effective interpreter feedback and instruction using GoREACT.com. Biennial Conference – Our Roots: The Essence of Our Future Conference Proceedings, 120 – 125.

Swan, S. L. (2015). Design thinking : New product development essentials from the pdma. Hoboken, New Jersey: John Wiley & Sons, Incorporated.

Tarricone, P. (2011). The taxonomy of metacognition. Retrieved from https://ebookcentral-proquest-com.ezp.oamk.fi:2047

Tyler, S. (2007). Approaches to learning and studying inventory (ASSIST) (3rd edition). Retrieved from ResearchGate: https://www.researchgate.net/publication/50390092_Approaches_to_learning_and_studying_inventory_ASSIST_3rd_edition

UNESCO. (2015). INTERNATIONAL STANDARD CLASSIFICATION OF EDUCATION Fields of education and training 2013 (ISCED-F 2013) – Detailed field descriptions. Montreal: UNESCO Institute for Statistics.

United Nations. (2020). Multilingualism in the United Nations system Report of the Joint Inspection Unit. Retrieved from https://www.unjiu.org/sites/www.unjiu.org/files/jiu_rep_2020_6_english.pdf

Universidad Pontífica Comillas. (2021). COMPETENCIAS ADQUIRIDAS. Retrieved from Comillas Universidad Pontífica: https://www.comillas.edu/postgrado/master-universitario-en-interpretacion-de-conferencias

Universidad Pontifica Comillas. (2021). Guia docente 2021 – 2022. Ficha tecnica de la asignatura Retrieved from https://repositorio.comillas.edu/xmlui/bitstream/handle/11531/57895/Gu%c3%ada%20Docente.pdf?sequence=-1&isAllowed=y

Universidad Pontífica Comillas. (2021). PRUEBAS DE ADMISIÓN MUIC. Retrieved from Comillas Universidad Pontífica: https://www.chs.comillas.edu/images/Documentos_K2/Estudios/PRUEBAS_DE_ADMISI%C3%93N_MUIC.pdf

Vandepitte, S. (2008). Remapping Translation Studies: Towards a Translation Studies Ontology. Meta, 569–588.

Walczyński, M. (2017). Assessing interpreting trainees' performance in a consecutive interpreting test with reference to affective factors: a proposal of a consecutive interpreting assessment tool. In J. M. (Ed), Cognitive approaches to specialist translation (pp. 121-135). Czestochowa: Publishing Office of Czestochowa University of Technology.

World Health Organization. (2020). Rehabilitation competency framework.

World Health Organization. (2021). Strategic plan 2021–2025: alliance for health policy and systems research.

Xu, R. (2018). Corpus-based terminological preparation for simultaneous interpreting. Interpreting, 33 – 62.

## 8. Acknowledgements

The author would like to thank the efforts of all those who made this research possible, helped in carrying it out and in making this process efficient and trustworthy (in alphabetical order):

Adriana Vieira, Anna Vegera, Blair Stevenson, Catarina Correia, Cristina Fontoura, Glenda de la Fuente, Helena Eijsberg, Janne Karjalainen, Johanna Pihlajamaa, Josh Goldsmith, Kenny Tran, Marisa Shirasuna, Marly Moro, Pablo Santur, Rosana Gouveia, Sieni Campos.

# 9. Appendices

Appendix 1: Invitation Template (type of meeting [interview/focus group] and timing modified as needed)

"PEER FEEDBACK IS KEY —
THE @IBPGS HAS DEVELOPED A FEEDBACK SHEET USED AT ALL SESSIONS,
AND MAKES THE PEER FEEDBACK EASY TO ACCEPT"
— says *AIIC Interpreters* on Twitter

By using feedback forms "TO TRACK YOUR FEEDBACK OVER TIME,
YOU CAN CREATE A FEEDBACK PORTFOLIO THAT GROWS WITH YOU AND REFLECTS THE SKILLS YOU ARE WORKING ON"
— adds *Techforword*, an online academy for language professionals

Numerous academic publications as well as syllabi emphasize the importance of feedback and assessment criteria

And, according to some practitioners of interpretation, finding the big picture and patterns of student performance can provide helpful insights into learning

**While the importance of giving feedback in training is widely recognized, the question remains: how?**

What do you think about giving feedback?

What do you think about the role of standardized patterns and measurable indicators?

How to keep feedback clear at all times?

I hope to find the answers in a series of one-to-one Zoom interviews with practitioners of interpretation

I am Mikhail Demidov,
a CELTA-certified teacher of English and a freelance medical interpreter / translator
with experience at various international organizations

I am currently working on **an app for interpreter training** as a part of my
Master's Degree in Education Entrepreneurship at Oulu University of Applied Sciences (OUAS)

I would like to invite you to take part in a study to gather your thoughts and experiences with giving / receiving feedback in the classroom

**It is my pleasure to invite you to join a 60-minute interview**
**comparing three prototypes of my app**

Participants will be provided a recording of a speech in English and its interpretation into a different EU language, about 5 minutes long. Each participant will be asked to assess the quality of the interpretation using three different prototypes of the assessment tool.

The meeting will be recorded for further automated transcription and qualitative data analysis. The findings will then be summarized in my thesis at OUAS. The validation with the professional community will serve the promotion of multidisciplinary skills and user-centered communication for the benefit of our industry.

**Join the meeting and contribute to an ongoing research into practical tools for interpreter training!**

## Appendix 2: FG consent form

**Consent Form for Participants**

Informed consent for participating in research

**Background**

My name is Mikhail Demidov, I am a freelance interpreter/translator with a working background with several UN agencies.

You are about to take part in a focus group that I conduct as a part of my portfolio-based thesis at Oulu University of Applied Sciences (OAMK) in Finland. At Oulu, I am working on tools for interpreter training. Presently, three prototypes of one such tool have been designed.

The focus group will help me create a peer-assessment tool. It is expected to help students of interpretation deal with some common challenges. Furthermore, successful iterations of the tool may be used in future to create a commercial product for interpreter training (e.g. a mobile app).

Before we move on, let me address some important arrangements

**Procedure**

I am going to ask you some general questions about your educational and professional background as well as your experiences with Translation and teaching (if you are currently teaching Translation or taught in the past; 'Translation' with a capital T to refer both to translation and interpretation as a cognitive process).

After that you will be provided a recording of a speech in English and its Translation into a different language (either your working language or your mother tongue), about 5 minutes long. You will then be asked to assess the quality of the Translation you hear using three different prototypes of the assessment tool. The description of the prototypes will not be provided in advance, as I am also going to explore if user interface of my tool is intuitive enough and user-friendly.

The focus group meeting will last about 120 minutes.

**Data collection**

The interview will be recorded for further automated transcription with Otter AI or NVivo and qualitative data analysis with QDA Miner Lite. The usage of video is optional. Only the audio from the interview will be used as data. I will not collect personally identifiable information, including your name, date of birth, phone number, e-mail address, or IP address. I also encourage you to use a nickname instead of your real name. You will be provided a link to the Zoom meeting in advance through your moderator, therefore, no prior exchange e-mail addresses or telephone numbers is necessary.

Please do not communicate any proprietary, confidential, identifying or sensitive information. Let's agree that any information that you (or persons acting on your behalf) provide in your responses is not proprietary, confidential, identifying or sensitive.

In the thesis, I am going to use solely the quotations from the transcripts to describe a variety of opinions about the skills and knowledge that interpreters need, as well as your experience with my prototypes.

I will destroy the recordings and other data after the thesis has been assessed and approved by OAMK. The thesis will be published in OAMK's Theseus database which may provide open access to readers. The

results of my research may be published in scientific journals, articles, conference papers, or in other media in future.

**Other provisions**

I would like to make sure that you understand and agree that your participation is voluntary, and does not create agency, legal partnership, or employment relationship. You may choose not to participate. You may also choose to withdraw from the research at any time. Information collected before your withdrawal will not be used. I'll still be grateful for your willingness to help in the first place.

Please feel free to contact me for any further information that you may need in relation with this interview.

***Mikhail Demidov***

**Confirming the informed consent**

I am willing to participate in the research.

I allow the use of my audio recording for research purposes.

I allow the information that I have provided to be used until the thesis has been assessed and approved by OAMK.

Date ___/___ 20___

Participant's Signature (signature, name)

_____

Researcher's Signature

_____

Mikhail Demidov m1demi00@students.oamk.fi +358465345259

**This thesis research is supervised by:**

Pablo Santur Arrelucé, Oulu University of Applied Sciences, Finland

**More information about research ethics and informed consent:**

Finnish Board on Research Integrity

http://www.tenk.fi/en/ethical-review-in-human-sciences

Social Sciences Data Archive

https://www.fsd.tuni.fi/en/services/data-management-guidelines/informing-research-participants/#partIV-examples-of%20informing-research-participants

http://www.fsd.uta.fi/aineistonhallinta/en/anonymisation-and-identifiers.htm

Appendix 3: Pre-meeting questionnaire

# A Survey on Interpretation Competence

Background information before the April 10 focus group

## Disclaimer (Notice to Survey Participants)

Hello! Thank you for your interest in my research!

I am Mikhail Demidov, a freelance medical interpreter/translator with experience at various international organizations, and a CELTA-certified teacher of English.

I am currently working on an app for interpreter training as a part of my Master's Degree in Education Entrepreneurship at Oulu University of Applied Sciences (OAMK), Finland.

You are about to take part in a survey that will ask you some general questions about the Theory of Translation and your teaching experiences (if you are currently teaching Translation or taught in the past; 'Translation' with a capital T to refer both to translation and interpretation as a cognitive process). The results of this survey will be described in my thesis alongside the focus group results. The descriptions will be generalized and will not lead to the identification of individual participants.

The survey will help me create a peer-assessment tool. It is expected to help students of interpretation deal with some common challenges. Furthermore, successful iterations of the tool may be used in future to create a commercial product for interpreter training (e.g. a mobile app).

Before we move on, let me address some important provisions.

First of all, I would like to make sure that you understand and agree that your participation is voluntary, and does not create agency, legal partnership, or employment relationship. You may choose not to participate. I'm still grateful for your willingness to help in the first place.

You will be asked a number of questions in Microsoft Forms. Most of the questions are about learners' skills and their relevance for interpreter training. The survey is not really long and should take around 10 minutes of your time. It will also help us speed up our focus group discussions.

A downside of this questionnaire is probably that Microsoft forms will not remember where you stopped if you happen to close the window halfway through the process.

Please do not distribute, sublicense, or create derivative works of this questionnaire.

I also would like you to know that I do not collect any identifying information, including your name, date of birth, phone number, e-mail address, or IP address. The survey does not contain information that personally identifies you. Data is stored in a password-protected Microsoft repository offered to OAMK students under a student account.

In order to minimize the risks of bugs and virus threats, I am using licensed software and antivirus tools.

The survey contains several free-form responses. Please do not to include any proprietary, confidential, identifying or sensitive information. Let's agree that any information that you (or persons acting on your behalf} provide in your responses is not proprietary, confidential, identifying or sensitive.

Please decide on further course of action

**By choosing the 'Accept' option, you certify that you have read, understood and agreed to the above. Otherwise please choose the 'Decline'** option or **close** this **window.**

Mikhail Demidov *

● Accept

● Decline

Below is a simplified schematic representation of Translation process as described by *Interpretive Theory of Translation* (Lederer, Seleskovitch).

Do you think it reflects your approach to Translation?

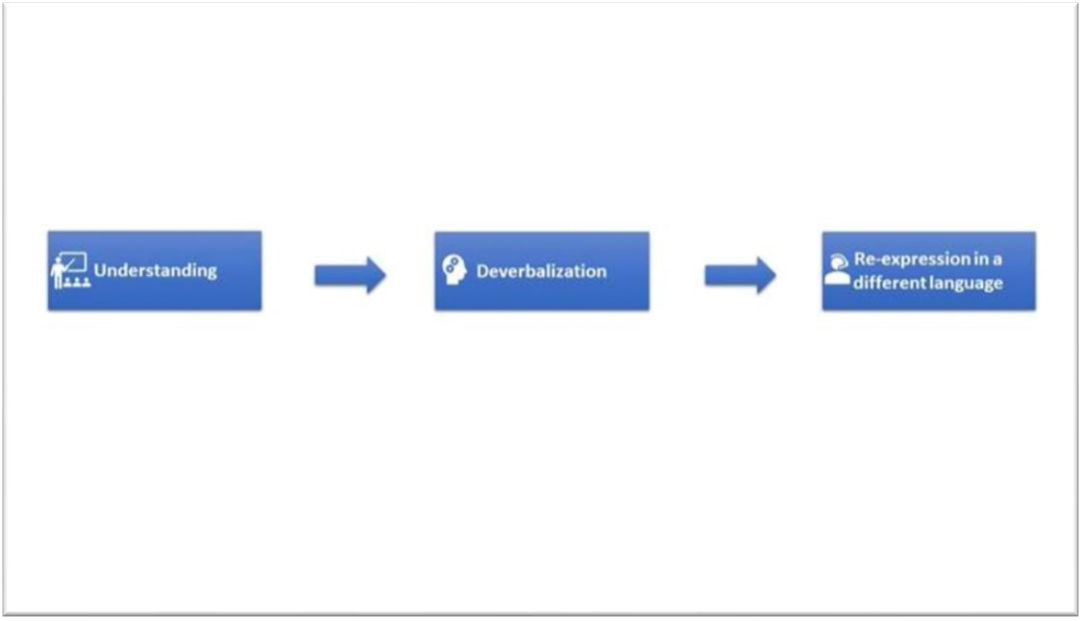
**Terms**

**Translation** (with upper-case T) is used to collectively refer to interpreting and translation as a cognitive process.

**Understanding** — Understanding a speech or a text is a process in which sense is extracted from an aural or graphic sequence through the combination of cognitive inputs with linguistic meanings.

**Deverbalization** — In the translation process, deverbalization is the phase that comes between the understanding of a text and its re-expression in another language. Verbal signs fade as cognitive and affective sense is grasped.

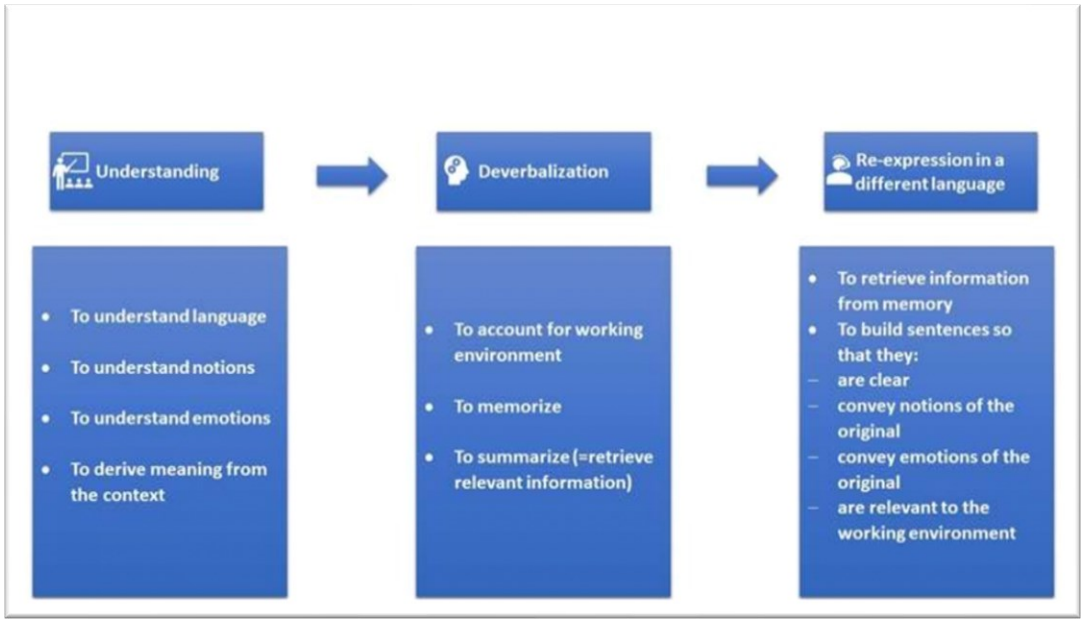
**Re-expression** — Finding a suitable formulation. A process based on the ideas understood, not on the words used to convey them



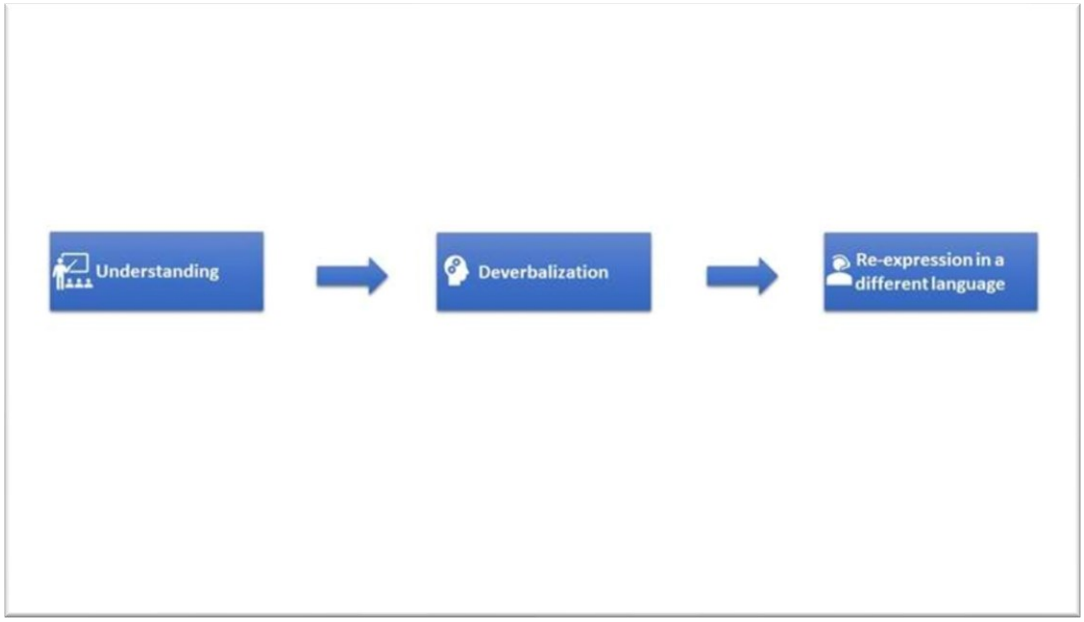

Source: Lederer, M. 2003. *Translation . The interpretive model.*

Below is the same diagram with a list of skills added that are probably relevant at each stage. Do you think it reflects your approach to Translation?

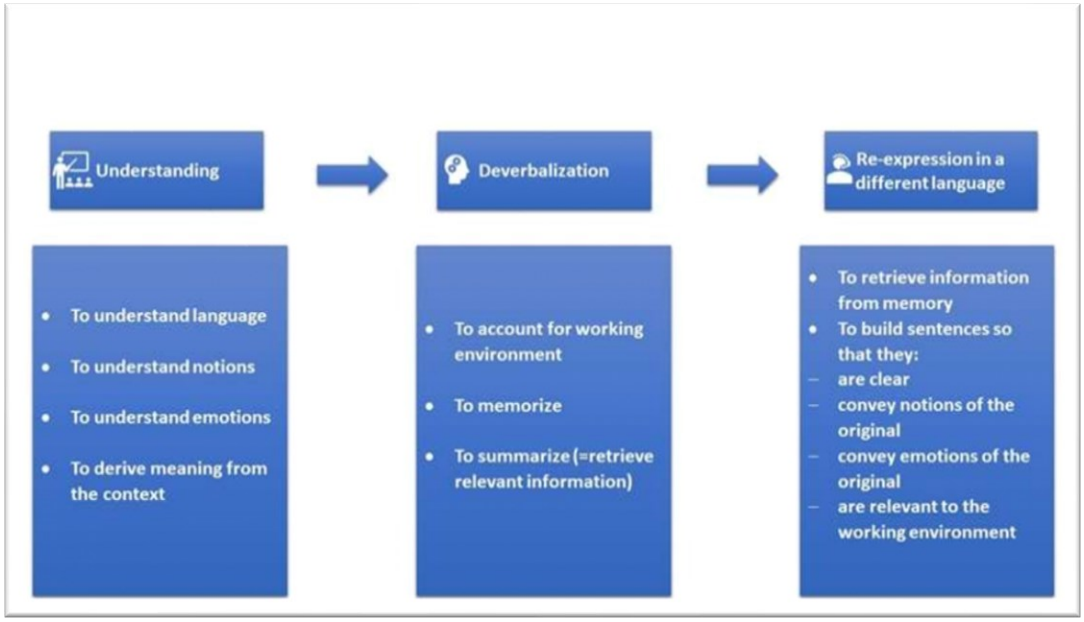
Note about the term 'Relevant information': As Patrick Winston put it in his lecture *Story Understanding,* relevant pieces of information are the ones that have links to other bits of information in the text, or lead to building new concepts.




If this diagram is not exactly reflective of your approach, could you clarify why?




If this diagram is not exactly reflective of your approach in terms of skills, could you clarify why?




This is a list of phrases ('descriptors') that you can use to characterize your challenges in Translation. Score them depending on their relevance to you (#5 = highest score, most relevant).

| | 5 | 4 | 3 | 2 | 1 | 0 |
|---|---|---|---|---|---|---|
| Tongue-tied | ○ | ○ | ○ | ○ | ○ | ○ |
| I wish no one listened / I wish it ended already | ○ | ○ | ○ | ○ | ○ | ○ |
| I can't work for so long! | ○ | ○ | ○ | ○ | ○ | ○ |
| I know all the words but they don't make sense | ○ | ○ | ○ | ○ | ○ | ○ |
| I see what the speaker means, but I cannot put it into words | ○ | ○ | ○ | ○ | ○ | ○ |
| The right word escapes me | ○ | ○ | ○ | ○ | ○ | ○ |
| I can't start the sentence | ○ | ○ | ○ | ○ | ○ | ○ |
| A bit too much detail | ○ | ○ | ○ | ○ | ○ | ○ |
| The longer I work, the more carried away I get | ○ | ○ | ○ | ○ | ○ | ○ |

Do you currently teach interpretation (or did you teach interpretation in the past)? *

Yes

No

## Questions for teachers of interpretation

Do learners achieve a learning plateau during the course? Is it common?

How many trainees would usually achieve the plateau in an average year?

Could you grade the skills below based on how quickly the plateau is achieved with each of them?(on a scale of 0 to 5; #5 = plateau is achieved quickest).

| | 5 | 4 | 3 | 2 | 1 | 0 |
|---|---|---|---|---|---|---|
| The usage of advanced oral presentation techniques | ○ | ○ | ○ | ○ | ○ | ○ |
| The usage of different coping strategies (processing / summarising speech) | ○ | ○ | ○ | ○ | ○ | ○ |
| The reproduction of the structure, style and register of the speech in question | ○ | ○ | ○ | ○ | ○ | ○ |
| The usage of different types of memory | ○ | ○ | ○ | ○ | ○ | ○ |
| The adaptation to the working environment | ○ | ○ | ○ | ○ | ○ | ○ |
| The usage of clear sentences | ○ | ○ | ○ | ○ | ○ | ○ |
| The usage of sentences that convey the notions of the original | ○ | ○ | ○ | ○ | ○ | ○ |
| The usage of sentences that convey the emotions of the original | ○ | ○ | ○ | ○ | ○ | ○ |

Do you think any important skills were missed in the above question?

This is a list of descriptors that you probably use to characterise student performance in Translation. Could you grade them in order of relevance to your work (on a scale of 0 to 5; #5 = most relevant)?

| | 5 | 4 | 3 | 2 | 1 | 0 |
|---|---|---|---|---|---|---|
| Sounds sophisticated | ○ | ○ | ○ | ○ | ○ | ○ |
| Doesn't sound confident | ○ | ○ | ○ | ○ | ○ | ○ |
| Odd intonations | ○ | ○ | ○ | ○ | ○ | ○ |
| Misinterpreted / contresens | ○ | ○ | ○ | ○ | ○ | ○ |
| Interpreter speaks longer than the speaker did | ○ | ○ | ○ | ○ | ○ | ○ |
| Uses wrong words | ○ | ○ | ○ | ○ | ○ | ○ |
| Makes their own additions | ○ | ○ | ○ | ○ | ○ | ○ |
| Makes omissions | ○ | ○ | ○ | ○ | ○ | ○ |
| Too much self-corrections | ○ | ○ | ○ | ○ | ○ | ○ |

Would you like to add your own descriptors or modify the existing ones so that they reflect thesituation more precisely?

## Comments

This is the end of the Survey form. Please leave a comment or make a suggestion if you like.

Thank you for taking your time!

Appendix 4: FG coding frame with quotes

| **Research sub-question** | Question being asked | **Classes (Hierarchy level 1)** | *Rules of assigning classes* | **Categories (Hierarchy level 2)** | *Rules of assigning categories* | **Subcategories (Hierarchy level 3, concrete ideas about categories)** | *Examples (Quotes)* |
|---|---|---|---|---|---|---|---|
| **What competencies should the tool support?** | A. What would be the most important skills and/or competencies for an interpreter? Could you talk me through the main criteria that you think are the most relevant? | | *What the FG participant mentions* | | *The FG participant talks about practical know-hows* | **A.1.1.1. Language mastery** | I think that one should master both languages. Experience in leaving in native country, I think it's highly valid, but it's not mandatory. |
| | | | | | | | Two speakers agreed with this point without further additions. |
| | | | | | | | Love the language that you are mastering. Mastering the language is important of course, you have to keep on learning forever and ever. So whatever your languages are, you have to work on them, read in them, listen to news in them, look for newspapers, news in them, read the articles. |
| | | | | | | | Sometimes you're just watching the soap opera or a film [and you think to yourself]: how would I translate that in my language, or that expression is really interesting. |
| | | | | | | | I have a degree in engineering, civil engineering, but I became an interpreter like 22 years ago, and that has been my profession to date. <…> Field knowledge is important. I mean, if you have a different degree, or any other experience working in the field, I think that helps. We all know that. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | **A.1. Theoretical expectations** | *before the actual assessment exercise* | **A.1.1. Practical skills** | *that have been acquired through training* | **A.1.1.2. Thematic knowledge** | I graduated in history of art and I hold a Master's degree in history of art as well, which I've taught for 11 years <…> and I've been an interpreter since 1986 <…> Topics are art, of course, social and human sciences and environment. |
| | | | | | | | Most of my work has been in international cooperation and business. And those are actually my favorite subjects. And also a lot of like all of us climate change, and now the pandemic and human rights and equity and gender issues, etc. I also do some medical, but it's not my favorite, and my hate subject is law. I really, really dislike working with law. |
| | | | | | | | I would work in other areas as well except for Law and nuts and bolts. |
| | | | | | | | I tend to agree with what everyone said. But to put it in order of priorities for me, kind of, I would say the most important thing is enjoying the work what you do the learning, because you're always plunged into different fields. |
| | | | | | | **A.1.1.3. Professional demeanour** | |
| | | | | | | | The second requirement, I'd say it's the ability to put ourselves in other people's shoes in order to understand that their discourse is coming from somewhere else, not from my point of view. |
| | | | | | | | I try to take myself out of the picture and think that the person listening to me should be understanding the message that's been given as faithfully as possible to what's been given, rather than me trying to sit and figure out what is the best word or you know, not so self-centered and more to think about the person who's listening to me. |
| | | | | | | | And what I try to do most of the time is, like Amanda said, I tried to take myself out of the picture and think that the person listening to me should be understanding the message that's been given as faithfully as possible to what's been given, rather than me trying to sit and figure out what is the best word or you know, not so self-centered and more |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | to think about the person who's listening to me, sometimes that's hard when the subject can be quite complex. And then you focus too much on your own performance. But that's what I work on most of my time. |
| | | | | | | | And I think it's really important to keep a high level of inner silence of being centered in yourself, so that you're not swayed by some word or some expression that you haven't understood or the audio that's really bad. And, you know, keep that inner silence as a compass in whatever you're doing. |
| | | | | | | | If you can, we try to just mime in the style the person is using, so that we can really produce the same effect on the audience, so listening to the interpretation feels the same way to the audience as listening to the original audio, and the same reacting. So I think that this thing of putting yourself in the other person's shoes is very important. |
| | | | | | | | So I think that we have to be very good pressure managers, because we suffer pressure from all sides, not only at the booth, but also when we receive the material to study and that usually is the midnight and your your event starts seven in the morning. And then you still have to go through his slides and everything else. So we have to know how to manage time, and how to manage pressure very well. Time management is because when we receive the material, what how much time we should have for our own private life. |
| | | | | | | | Train yourself to have nerves of steel or at least pretend that you do; try not to crumble in the booth or with your colleague or especially if you're in a room full of people and you have a microphone in your hand, try not to shake that's important to give a good impression. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | **A.1.2. Personality traits** | *The FG participant talks about the patterns of thoughts, feelings and behaviours* | **A.1.2.1. Being determined** | To be a little bit bold, because sometimes we get there and we're not fully ready. So one must be bold. I'm talking about, you know, personality features. |
| | | | | | | | Let's say, a leap of faith, believing that you can do it at the moment the event starts because otherwise, you're kind of frozen into place and you can't do go forwards. |
| | | | | | | **A.1.2.2. Being mindful** | So I just, I just want to say that we also have to be more indulgent on ourselves, we tend to judge ourselves too much. And especially that word that we'll never forget, we never learn that specific word that we cannot remember. So I think that we have to learn how to, you know, forgive ourselves for the mistakes, we are not perfect, we are not people that are there just to be 100%. Okay, we know that this is implied in our job, but many times we just tend to, you know, just sacrifice ourselves for many things. But when we just do one little mistake, we tend just to be sad for all the rest of the day. |
| | | | | | | | And I would say I was gonna say to be kind to yourself, but Alice already said that, to be humble. |
| | | | | | *The FG participant talks about the* | **A.1.3.1. Willingness to learn and/or adapt** | I would just add flexibility. Because we really navigate in different settings. And that requires to be flexible. Many times, we don't have the ideal conditions. Yeah. And it's also different to behave in a setting where we are in a meeting with politicians, and you need to respect the protocol and be aware about these protocols. And when you are in the floor of a factory. So that requires a lot of flexibility. |
| | | | | | | | And if you don't enjoy the research and learning about new stuff, it's terrible, and no, you can't do your work properly. So you have to enjoy the dive. Because for what from what I've read, and from what I learned in practice, I think for every hour in the booth you have at least two hours outside |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | **A.1.3. Aptitudes** | *potential to acquire particular skills or competencies through training* | | the booth studying, if not more. So if you don't enjoy the groundwork, don't do this job, it doesn't work. |
| | | | | | | | And if you found something, piece of, of nut or bolt or a piece of law that you really don't understand, try and study it regardless of how tired you are, you have to be sorrow because it won't be sure. It's like school exams and it will come up. |
| | | | | | | | And just as being a learner, not only fast learner, and not only someone who has a passion for learning, but I think that we are, we have to be an eternal learner. |
| | | | | | | | So passion for learning, I think it's a must. |
| | | | | | | **A.1.3.2. Being open to new developments** | There's always some unknown. And we never know what to expect. We never know when we might come across a person whose line of thinking is not so clear. Or maybe we did not get aligned with the person or with the topic. So we always have to be ready for challenges. And this is part of our daily lives, and we are never ever ready. |
| | | | | | | | You have to think on your feet, be very, very flexible. |
| | | | | | | | And the one more aspect would be fast reasoning. Think fast, because what you are hearing might not be what you expect to hear. |
| | | | | | | | I think also that we have to learn how to incorporate new terminology very fast. Sometimes we are just learning all new vocabulary as we go through all these slides of the of the speaker, right there on site and right there at the moment of the interpretation. |
| | | | | | | | While I understood that the staggered was somebody who was doing simultaneous, and then there was like, patched up in order for it to sound consecutive. Oh, and that made it really hard for me, because I felt that I'm not really evaluating a real situation. So I was feeling kind of like, what's the purpose of this? You know, I'd rather either listen to both of them, at the same time, the way we do when |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | **B.1. Practical observations** | *What the FG participant mentions after the actual assessment exercise* | **B.1.1. Appeal to the assessor's mental model** | *The FG participant mentions the difference or some discrepancy between the "real situation" and the assessment situation, or between how they see the situation and how the tool suggests to see the situation* | **B.1.1.1. The assessment situation should reproduce the real-life situation** | we're in the booth or have a real consecutive interpretation, because I do not think that you can really take similar changes and transform it like this into consecutive. |
| | | | | | | | I listed fluency expertise, completeness, delivery, speed of delivery and intonation as criteria. But as I said, I did not think they applied to the questionnaire sound or I couldn't apply them to the spider thing that you sent either because as I said, the feeling I had is that this was there was a mismatch. And so was not happy either with the criteria of what we've what I heard, as I said, the feeling I have is this was a simultaneous interpretation with someone reading, the other one trying to follow and I didn't think the criteria for staggered or for consecutive interpreting applied. |
| | | | | | | | I had a lot of trouble. And I felt I wasn't able to evaluate it. Because actually, I was seeing it as a simultaneous interpretation in two parts. I wasn't able to see it as a consecutive. And it wasn't a real world situation in which we assessed the consecutive. |
| | | | | | | | I don't actually see the point, why would you be interested in if the what we heard was relaxing or accurate? Because we have to translate whatever comes your way. Okay. So this is something we will not have control over in real situation. |
| | | | | | | | When I was reading, you know, through the questions. I thought the same as I thought, well, this is about our early conversation when we spoke about the parameters and the criteria that we would be assessing on. So I thought, Well, that's about this. This is not about the lecture that I just |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | **B.1.1.2. The assessment tool should strike a chord with the assessor's mental model** | heard, because we saw the words like "formal" or "engaging". So it was unrelated. There was a mismatch. |
| | | | | | | | I think that the evaluation was more regarding the language elements. Instead, we mentioned many things that have to do with personality and all that [which is why I included a different list of criteria]. |
| | | | | | | | I kind of answer, but it's sort of random, because I don't even think it applied to what I heard. |
| | | | | | | | And for me, it was not what I expected for the form that you sent. I saw it in other eyes. And that's what I told you. For me. There was a discrepancy and all that. Not much more to add. |
| | | | | **B.1.2. Relative weight of criteria and the assessment context** | *The FG participant talks about why an issue occurs, asks why assess something in the first place, or whether some criterion* | **B.1.2.1. Relative weight of criteria** | And the guy was trying to follow up. And of course, he had to drop some sentences. And of course, some things were left out because there was no other way. That's my reading. But as for completeness, expertise, I think he tried to give some intonation where she didn't give much of any intonation. At some points, he tried to give some more. He presented his tone of voice he tried to give it some intonation, but she didn't, she just read on. That's my particular feeling. |
| | | | | | | | And one of the things that I do when I interpret and I also do that when I evaluate a student. First, what kind of speech do we have? This is just basically information that is necessary. This is not a motivation, conference, this is not something that we … your style will make any difference. And if especially if this is something that just comes up at a time, that is just a sudden, something that came up, you do have to interpret. Okay, so then second step, who can focus on style. So if the person is trying to be funny, if the person is trying to give stresses on some places or the other in a sentence, that's fine, then you try to do that. But if this is |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | *is really informative* | | something that you're not really prepared, is this a topic you're not do not really matter so much, technology can be a problem to look at. So first, focus on content, focus on trying to be as accurate as you can forget all the rest. |
| | | | | | | | [The interpreter on the recording missed some pieces] But the thing is, you cannot really blame the interpreter for that. We don't know how … what kind of preparation he had, maybe this just wasn't something that came up on time, we know that many times our speeches come up like this, and you don't really have any, any idea of what the person is going to say. And when the person sticks to reading, that's really something very hard for us to try to keep track on. |
| | | | | | | | I would say that the speaker who was speaking in English was reading a text, who was not get that that was not given on to the interpreter who was speaking Spanish. English is like, much shorter than Spanish, Spanish is around 20 to 25% longer when you have to translate than English. |
| | | | | | | | In a situation like that, you have to be … to try to be more precise contentwise, you have to be more adherent to the content because every information there is important, the way the person is saying the words, and this is also true when you have an audit or a consultancy, when you have a deposition when we are listening to a witness, every word including the one the person hesitates to speaking is important and this is what the attorneys on the other party are really evaluating, assessing. So you have to consider different situations. There are … well there are … there are times when you can be more concise, focused. There are other times when I'm just getting the sense, a gist of what the person say is enough. And there are different times when the more literal you can be is important. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | **B.1.2.2. Assessment context** | Yes, I find it was a bit. I found it that difficult to choose what to look for, at first, I was a bit lost. So what should I do? You said, okay, just choose some criteria and go on with that. Okay, me at that time, I think I got a little bit lost, and I didn't know what to look for. And then on the go, I was saying, Oh, I'm going to look for I'm going to pay attention to this. Oh, no, no. Okay. Okay. So it was not very consistent. |
| | | | | | | | In my case, I anyway send it. So it wouldn't be that useful anyway, because I did it because I have to do it. And I have to say something in a way. Yeah. Although I felt that it didn't match or it didn't apply and all that. Anyway, send it so that we were evaluating someone assessing someone, it would be real be unfair. |
| | | | | | | | But when you are to actually assess an interpreter, as a trainer, you usually go to the more technical elements, more technical components. That's, that's what I felt. Right? |
| | | | | | | | And then the second part, we just studied the, what we heard on the tape. And as it was said, we became more technical. |
| | | | | | | | So these are two completely different things. So if this had been a simultaneous interpretation, I would have made a certain kind of evaluation for consecutive, I think it's different. So that's what got me confused. So I was listening with a different ear. |
| | | | | | | | Maybe if you were to play the same segment, again, because I took notes as if it were consecutive, I knew that you were doing the staggered simultaneous, but I took notes to see if he was you know, translating everything. But any worked, because I could do a good assessment. |
| | | | | | | | I started thinking about accuracy. But then they realize that they should see this more as a consecutive rather than an interpreting kind of job. And then I said, Okay, so I should |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | take notes. And then my mind got confused. And then I shifted. I said, Okay, let's think about the register. |
| | | | | **B.1.3. Technical skills (mentioned during the focus group AND included on the list in Prototype 3)** | *The FG participant talks about specific competences (knowledge, skills, attitudes) which enable the interpreter to meet the needs of their work. Usually these competences have one-word names and are often found in textbooks and guidelines* | **B.1.3.1. Fluency** | Okay, what happens is that I chose a completely different elements to assess upon doing the actual technical, simultaneous translation. What came to mind when I was assessing this interpreter was *fluency*, terminology accuracy, self-confidence. of confidence has to do with what we discussed before. |
| | | | | | | | So, yeah, I included assertiveness, coherence, *fluency*, register, accuracy, face content rendered. |
| | | | | | | **B.1.3.2. Terminological accuracy** | Okay, what happens is that I chose a completely different elements to assess upon doing the actual technical, simultaneous translation. What came to mind when I was assessing this interpreter was fluency, *terminology accuracy*, self-confidence. |
| | | | | | | | So, yeah, I included assertiveness, coherence fluency, register, *accuracy*, face content rendered. |
| | | | | | | **B.1.3.3. Coherence** | So, yeah, I included assertiveness, *coherence*, fluency, register, accuracy, face content rendered. |
| | | | | | | **B.1.3.4. Adherence to the original** | And I added adherence to original. |
| | | | | | | | So, yeah, I included assertiveness, coherence, fluency, register, accuracy, *face content rendered*. |
| | | | | | | **B.1.3.5. Register** | So, yeah, I included assertiveness, coherence, fluency, *register*, accuracy, face content rendered. |
| | | | | | | | I started thinking about accuracy. But then they realize that they should see this more as a consecutive rather than an interpreting kind of job. And then I said, Okay, so I should take notes. And then my mind got confused. And then I shifted. I said, Okay, let's think about the *register*. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | **B.1.3.6. Delivery** | The same here, I changed or I added a few criteria. And the most important, I think, is delivery. |
| | | | | | | | I was a bit lost about what criteria I should serve. But when I started listening, I changed the criteria, for instance, delivery. I was not among the criteria and so I added it. |
| | | | | | | **B.1.3.7. Completeness** | I listed fluency, expertise, *completeness*, delivery, speed of delivery and intonation as criteria. |
| | | | | | | | I also added expertise on the topic, *capability to recover from misinterpretation*, completeness. |
| | | | | | | | I liked the words that he chose to recover from mistakes. I didn't see any, any major mistakes. And *completeness*. This is what I listed. |
| | | | | | | **B.1.3.8. Ability to recover from mistakes** | I liked the words that he chose capability to *recover from mistakes*. I didn't see any, any major mistakes. And completeness. This is what I listed. |
| | | | | | | | I also added expertise on the topic, *capability to recover from misinterpretation*, completeness. |
| | | | | | | **B.1.3.9. Conciseness** | It was adherence to original delivery, fast reasoning, self-centeredness, conciseness, nerves of steel, fast learning and terminology. And I'd like to stress the *conciseness*. This is very important to me because this is how your listeners pay attention, not to what you are saying, but to what the speaker is saying and don't get distracted. |
| **What characteristics should the user interface have?** | C. Could you talk me through your experiences with the user interface? | | | | *The FG participant* | **C.1.1.1. Freedom for** | The Excel form? Yes. I was a bit lost about what criteria I should serve. But when I started listening, I changed the criteria, for instance, delivery. It was not among the criteria and so I added it, but well, it [blank Excel form] was easier than the second one [pre-filled MS Forms questionnaire]. |
| | | | | | | | When I compare the first, the first form for us to evaluate to the second one, I thought that the first one gives us much more control, we, because we are going to choose, we're |

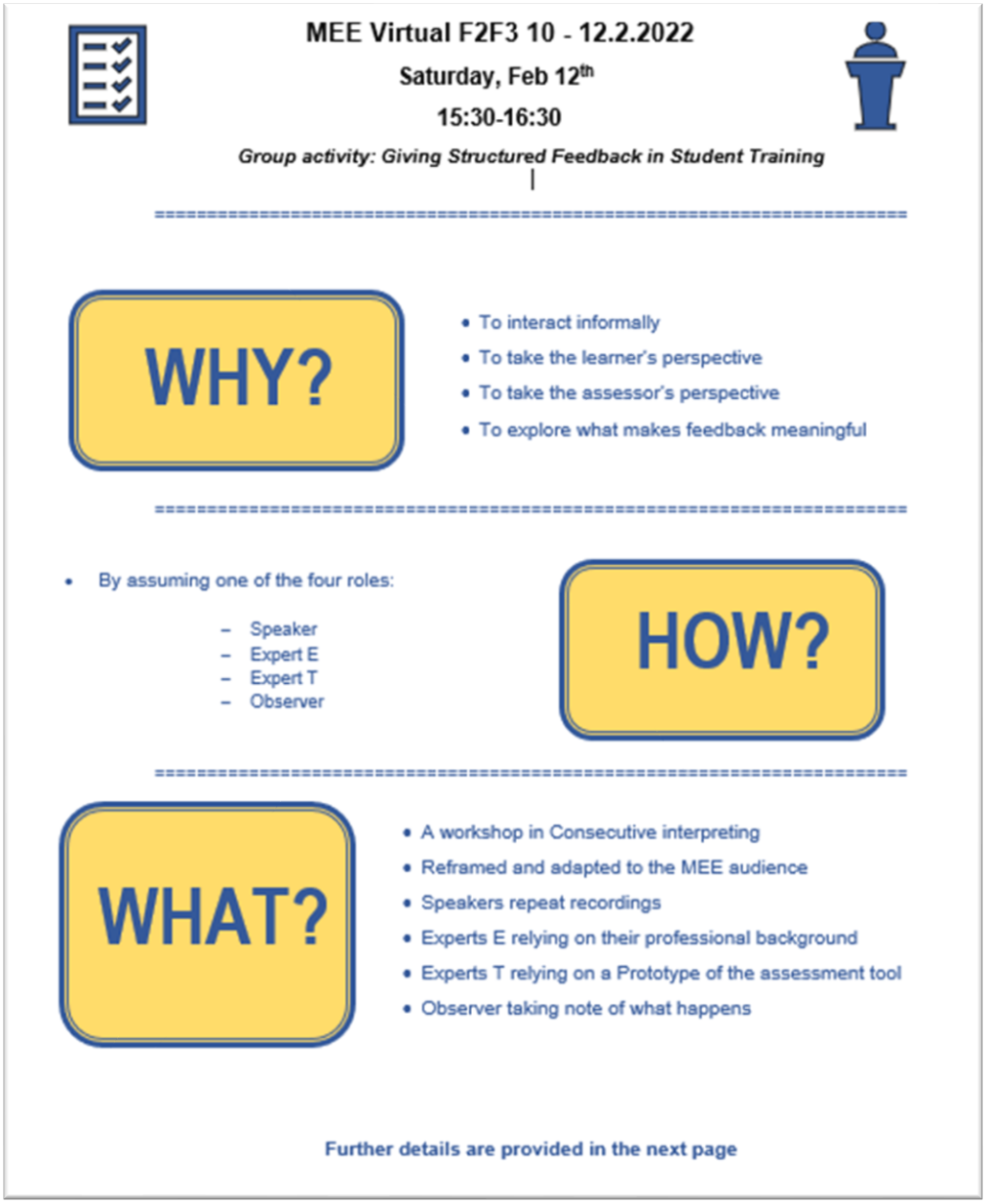
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | **C.1. UI characteristics** | *What the FG participant says about UI after the actual assessment exercise* | **C.1.1. The ease of handling the tool** | *expresses a positive or a negative attitude toward the elements of the user interface* | **user to determine criteria** | going to choose what we're going to look at. And the other. The other form is like a one size fits all. And then it didn't fit exactly that particular speech. And so this is how I see the differences between the forms. |
| | | | | | | | I fully agree with her. It when we have more freedom to identify the components that we are willing to assess. It's much better, because the second form Yeah, it was mind boggling. We were all wondering what is this about because it didn't fit it a match. So and it was preset, so perhaps it's better when you can develop your own assessment criteria. |
| | | | | | | **C.1.1.2. Saving the results** | Oh, by the way, I can't even save it [Prototype 1] also. So I'm completely at loss, I must say. |
| | | | | | | **C.1.1.3. Making the tool more self-explanatory** | I think that for pedagogical reasons, it would have been useful just to go over the form, instead of just giving it to us and say, Okay, this is it. We were really not prepared to what we were, what would we have to do? Yeah, then we realized that perhaps it didn't match quite well, the kind of setting we are given now |
| | | | | **C.1.2. The relevance of pre-filled prompts** | *The FG participant specifically addresses the prompts offered by the tool* | | Let me see. Question number two “They were crystal clear, you didn’t have to listen twice”. Now, I wouldn't have to listen twice if this were real interpretation situation. But then “They sounded right for this meeting. Neither too academic nor too informal” doesn't apply. “They sounded relaxed and natural, just like a story” doesn't apply. Also, that's number five [All sentences were complete. Speakers did not drop phrases]. That would apply. Number six, yes [The story sounded logical and coherent]. Number seven [Names and numbers were credible and realistic] would apply. Number eight does not apply [The speakers were engaging]. And the number nine doesn't apply [The speakers were keeping it short]. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | [I was looking at the assessment criteria in the form] and I was thinking about the conversation we had during the session, I was not sure that you refer to the speech. Actually, I was a bit confused. A bit confused. |
| | | | | **C.1.3. Features to add** | *The FG participant talks about some features that they would like to see* | | I was thinking about just having an assessment in a way of an application. So the importance of defining what you mean by that and also the possibility of leaving, like, it does not apply. Like, I don't need to, for instance, mark any stars, if it is true stars, so it may not apply and perhaps include that as part of the script before starting the evaluation, something like that. So in a way, people are all on the same page. Yeah. So we have a better understanding of it or have a possibility not to attend, just leave it with zero stars. Yeah, because it does not apply as an option. |

Appendix 5: List of criteria suggested by the FG participants (saved in Prototype 3)

1. Accuracy (mentioned 4 times)
2. Accurate vocabulary
3. Adherence to original
4. Articulate properly – avoid mumbling
5. Avoid empty words, i.e. ‘ah’, ‘oh’, ‘no’, etc.
6. Coherence (3 times)
7. Completedness
8. Conciseness
9. Confidence
10. Content (2 times)
11. Content
12. Delivery
13. Faithful to speaker
14. Fast thinking
15. Fluency
16. For fast speaker, keep the content, drop un-essentials
17. For slow speaker deliver as close to 100% as possible
18. Include names and numbers
19. Intonation
20. “Mime” the speaker as much as possible
21. Prior preparation
22. Register (2 times)
23. Self-confidence
24. Terminology
25. Tone of voice
26. Use of expert terminology
27. Use of formality or informality
28. Use of specific terms
29. Use of specific vocabulary

Appendix 6: Invitation template for the MEE workshop



**MEE Virtual F2F3 10 - 12.2.2022**

**Saturday, Feb 12th**

**15:30-16:30**

***Group activity: Giving Structured Feedback in Student Training***

==========

**WHY?**

- To interact informally
- To take the learner's perspective
- To take the assessor's perspective
- To explore what makes feedback meaningful

==========

**HOW?**

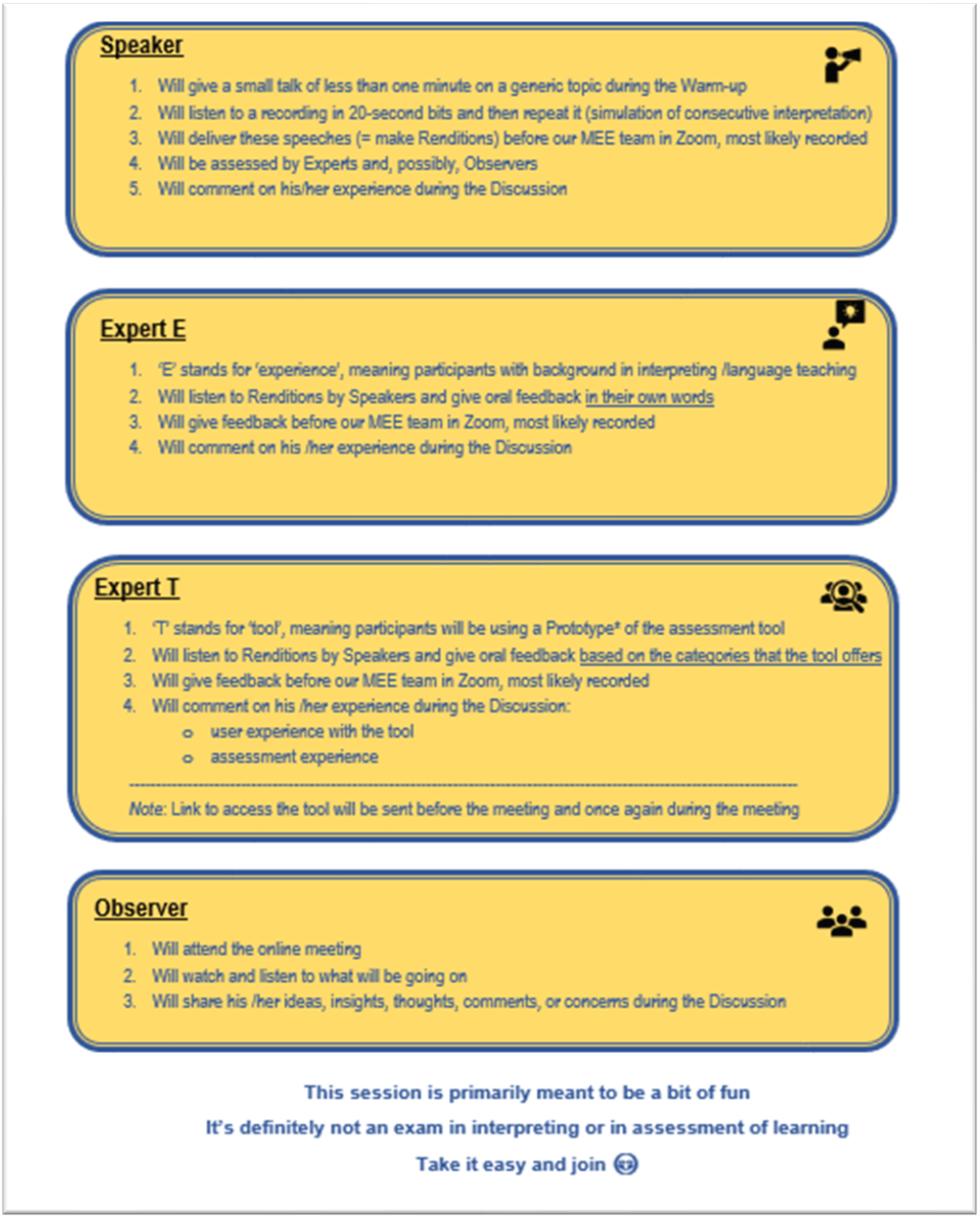
- By assuming one of the four roles:
  - Speaker
  - Expert E
  - Expert T
  - Observer

==========

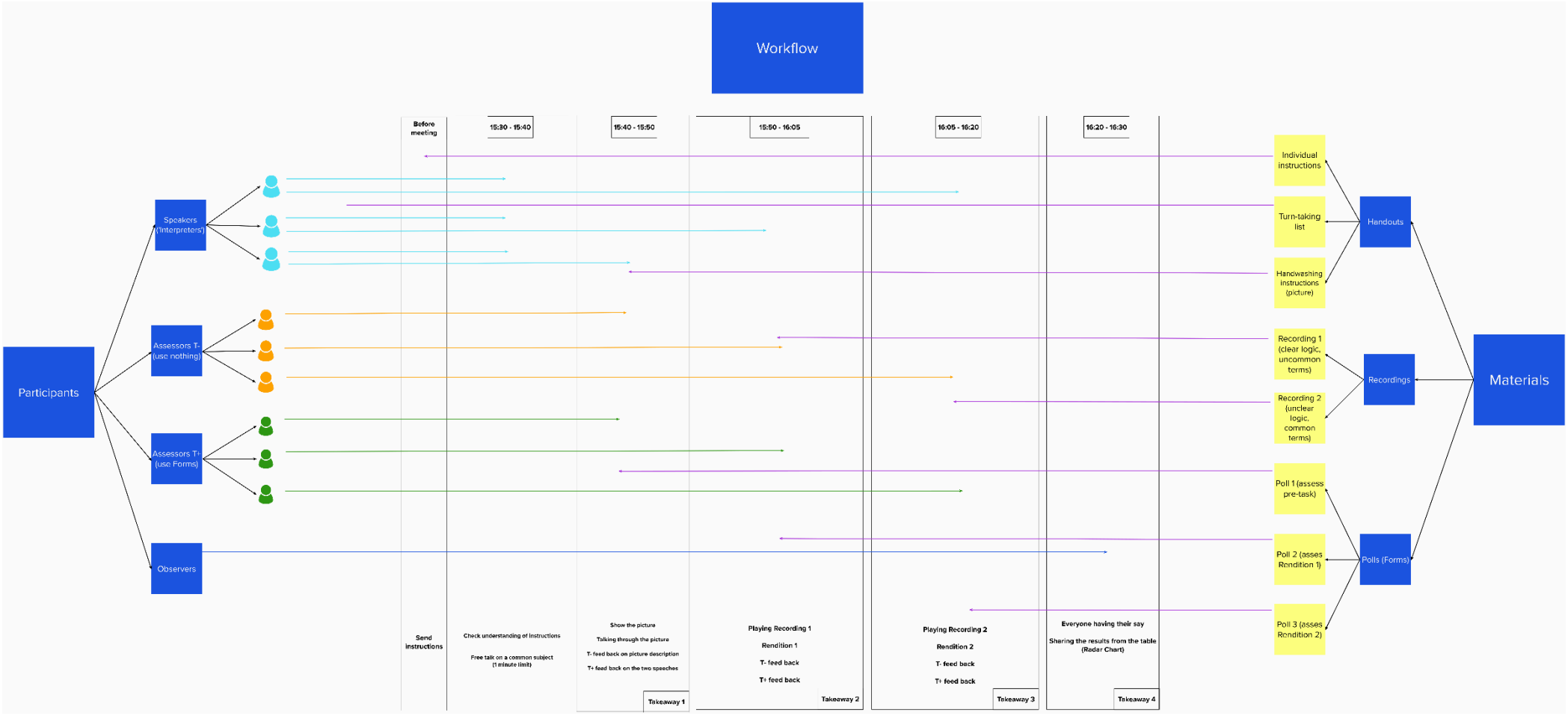
**WHAT?**

- A workshop in Consecutive interpreting
- Reframed and adapted to the MEE audience
- Speakers repeat recordings
- Experts E relying on their professional background
- Experts T relying on a Prototype of the assessment tool
- Observer taking note of what happens

**Further details are provided in the next page**


Speaker
1. Will give a small talk of less than one minute on a generic topic during the Warm-up
2. Will listen to a recording in 20-second bits and then repeat it (simulation of consecutive interpretation)
3. Will deliver these speeches (= make Renditions) before our MEE team in Zoom, most likely recorded
4. Will be assessed by Experts and, possibly, Observers
5. Will comment on his/her experience during the Discussion
Expert E
1. 'E' stands for 'experience', meaning participants with background in interpreting /language teaching
2. Will listen to Renditions by Speakers and give oral feedback in their own words
3. Will give feedback before our MEE team in Zoom, most likely recorded
4. Will comment on his /her experience during the Discussion
Expert T
1. 'T' stands for 'tool', meaning participants will be using a Prototype* of the assessment tool
2. Will listen to Renditions by Speakers and give oral feedback based on the categories that the tool offers
3. Will give feedback before our MEE team in Zoom, most likely recorded
4. Will comment on his /her experience during the Discussion:
o user experience with the tool
o assessment experience
Note: Link to access the tool will be sent before the meeting and once again during the meeting
Observer
1. Will attend the online meeting
2. Will watch and listen to what will be going on
3. Will share his /her ideas, insights, thoughts, comments, or concerns during the Discussion
This session is primarily meant to be a bit of fun
It's definitely not an exam in interpreting or in assessment of learning
Take it easy and join

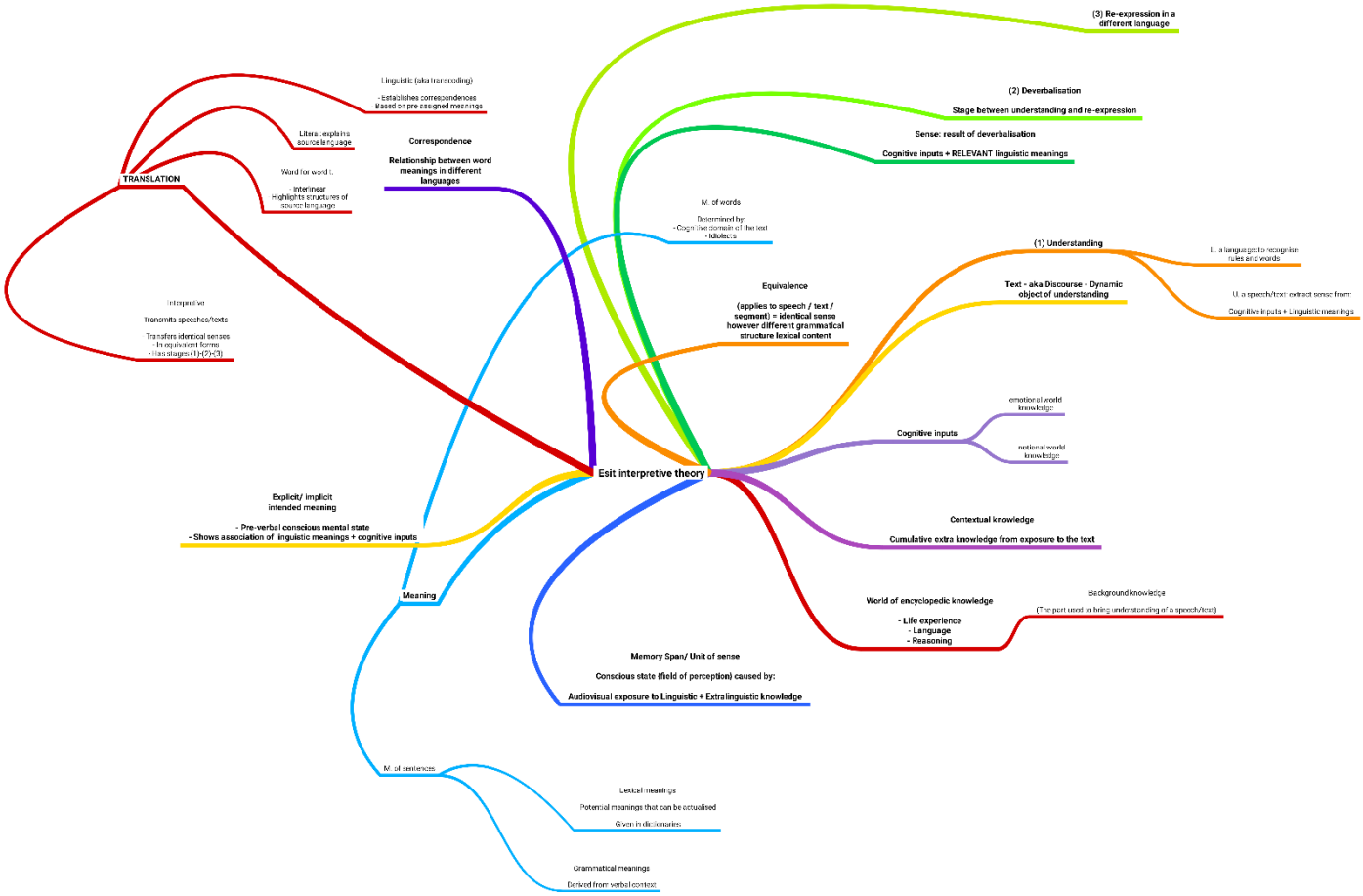
Appendix 7: MEE workshop workflow



# Appendix 8: MEE workshop coding frame with quotes

| Research questions | Concept-driven categories | Data-driven codes | Rules of assigning codes | Data-driven subclasses | Rules of assigning subclasses (mostly 'indicator' criteria, i.e. explanations contain specific wordings) | Examples (quotes) |
|---|---|---|---|---|---|---|
| What competencies should the tool support | 1. Self-reported competencies | 1.1. Immediately reported | *The competencies that a person says have helped them complete the task. The first thing that springs to his/her mind when asked immediately after their rendition. A question may be: 'what was your strategy?', 'what did you do?', 'what did you feel like? why?', etc.* | 1.1.1. Targets of communication | *The person addresses the circumstances / wider context /background of the speech in question.* | I feel very, I felt very, okay with this. Comfortable. I did feel the need to explain it a bit more. Because you asked me to speak for about one minute and I could probably do this in 10 seconds if I wanted to. So that's why the context was there. |
| | | | | 1.1.2. Consistent message | *The person mentions "sense", "making sense", "getting the message through", etc.* | Yes, I am not good at repeating things immediately. And I was trying to focus on having all the words, but like, at the beginning, I was trying to have all the words, but then I just switched to listening first to what the woman was saying, and really understanding what she was saying. And then it was easier. Because first I was just focusing on on memorizing the order of the words, and I wasn't really listening to the context. But then in the end, I kind of got the hang of it. And I was just Yeah, I'm just trying to give back the meaning. Or at least make a little sense in my repetition. |
| | | 1.2. Reported on second thoughts | *The competencies that a person says have helped them complete the task, when asked some time after their rendition, and other participants have had their say. A question may be: 'in what way was your strategy different from other participants?', 'is there anything else that you can think of when reflecting on your rendition?', etc.* | 1.2.1. Familiarity of the topic | *The person mentions being good or bad at some topic, having good or bad command of some topic, being familiar or unfamiliar with some topic.* | Um, I don't know. It's a more familiar topic. Like for me, when I just read the title, like the scientific use of eggs, I immediately was thinking in my head, I'm like, Oh, my God, what am I going to be listening to? But then, like, when I saw K's title, I was like, Oh, he got the easier one. |
| | | | | 1.2.2. Usage of backup strategies | *The person mentions having a difficulty and overcoming it due to a specific technique (such as taking notes, memorizing in a special way, etc.)* | I had to look what's on the screen text, take what I can see to interpret the process, I guess stages from 1 to 5. |
| | | | | | | Yeah, I followed S's experience and decided not to take notes and focus really on like, what do I hear? Do I understand what I hear? And then I quickly realized that I clearly understand what I hear. But it's still very difficult to remember six languages. So I kind of made up a few of them on the go. |
| | 2. Observer-reported competencies | 2.1. Immediately reported | *The competencies that a person says have helped someone else complete the task. The first thing that springs to his/her mind when asked immediately after the rendition. A question may be: 'how would you characterize X's performance?', 'what did X exactly do?', 'what would you feel like if you were offered the same assignment as X? why?', etc.* | 2.1.1. Being clear | *The person characterizes someone else's rendition as "clear", "easy to understand", etc.* | I think there would have been able to understand what definitely to understand what it was about. |
| | | | | | | Okay, I think she did a great job. I mean, just like, like M said, I feel that most of the time, she repeated what she heard. Her reaction was really positive, because she understood. I mean, what the lady on the screen was, was saying, so I think she did a great job. And I understood like, everything. |
| | | | | 2.1.2. Reproducing or losing detail | *The person says someone else have rendered all the details or lost some of the details. They usually give examples of such details.* | But sometimes maybe I will feel I feel a bit confused, because they will kind of kind of hesitation when it comes to certain figures and numbers. But I will understand, I will definitely understand what is the topic about and I mean, the whole thing, but they will be kind of like, maybe five to 10% of the confusion. |
| | | | | | | But of course, in the middle of the process, she forgot some things. |
| | | | | | | But of course, in general, I think like 70%, he understood well and did his best. |
| | | 2.2. Reported on second thoughts | *The competencies that a person says have helped someone else complete the task, when asked some time after the rendition, and other participants have had their say. A question may be: 'in what way was X's strategy different from other participants?', 'is there anything else that you can think of when reflecting on X's rendition?', etc.* | 2.2.1. "Overcoming "technical" difficulties of the speech | *The person mentiones some auxiliary techniques that might have improved someone else's preformance.* | Well, I think she did well, because there were a lot of adjectives and very specific words, and she remembered a lot of them. But it was taxing memory because of so many, perhaps adjectives and things and details. Yeah. But she kept the like the semantic units. So she made sense of what she was saying. Yeah, although she forgot some details. Overall, it made sense. She was able to just to, to pass on the message. |
| | | | | | | Quite well, I mean, he did his best but of course maybe taking some notes in this particular case would be like, a little bit better. |
| | | | | | | I think that in our render, like, in the same pace, which was good, yeah. And well, he he got wrong. A couple of things regarding accuracy, you have to change some names. Like in those days, I agree that taking down notes would have helped. |
| | | | | 2.2.2. Competencies that lacked | *The person usually gives feedback using "but" sentences. The first part of the sentence praises an achievement, whereas the second part reveals missing (and, obviously, important) competencies.* | — If you put yourself into S's shoes, what would you feel like?<br>— An immense pressure because when you listen, it's doable, but it seems to be a lot of information in 10 seconds. |
| What is the user experience with the tool | 3. Experience giving feedback | 3.1. Advantages | *The person makes a reference to any features of the tool or any aspects of the workflow that have reportedly improved their feedback or helped them with their feedback.* | 3.1.1. A feature which is helpful in general | *The person speficy a feature or an aspect of the workflow which is "nice", "interesting", "helpful", "complementary", etc.* | But in the third or fourth time I understood the how to do it, and it solved the rapidly. The last stuff I want to say is, so I loved that tool, because the more you use it, and the more you are capable in using it, the last stuff lost a thing is related to something I'm seeing into the answers. |
| | | | | | | Like, you know, the, the informal feedback that I received afterwards was indeed quite nice. And this one is a bit more depth, like it says, it's a bit strange to see also at a reset the red line going in. So that's definitely something that I would focus on very much like, what happened there? Like why, why is that scored like this? And how can I make that line move outwards? Again, that would be my my key takeaway from this chart. But I think from a personal point of view, I definitely consider myself a positively focused person. So I am also really happy to see a lot of the on the right and the bottom side, there's a lot of - how do you say that - everything's is scored quite nicely there. I think that red line can also do me justice, because that could be a fair assessment. |
| | | | | | | Well, it's definitely this feedback definitely gives me a lot more information. I think if I compare it to the unstructured feedback that I got right away the oral feedback, I'm like, if I compare this to that, this is this gives me a lot more information. |
| | | | | | | I see. Like, I don't know, if I say like, parameters, indicators, KPIs and how to say that. But of course, if we will have this before the session starts, of course, I will feel like better, like, after seeing this chart. Now. I think that maybe I missed some points. But of course, this is the exercise. |
| | | | | | | So I definitely think that having this kind of feedback in the making the chart, it helps me more in the sense that I can I always have that so I can even like hours later, if I get home from the conference or whatever. I can sit down and actually look through the chart. And while if I just get immediate oral feedback, I'm might forget it. Like maybe there was something very important that was said, but people and I don't usually tend to note down feedback. So I would probably forget about it. And then in on the other hand in the chart, maybe I can also see my progress through weeks if I'm if I'm evaluated and constantly with this |
| | | 3.2. Downsides | *The person makes a reference to any features of the tool or any aspects of the workflow that have reportedly complicated or invalidated their feedback.* | 3.2.1. A minus, but not a barrier | *The person thinks that they had difficulties giving feedback because of some feature of the tool, or the tool has not allowed them to make feedback complete. However, they do not think it invalidates their feedback.* | I would have loved just to write some some personal comments on it's just too because now I know, for example, that I gave K a three, but I don't really know anymore why I gave him a three or what my, my thinking was when I gave him a tree. |
| | | | | | | What it is relevant, yes, because it's a structure and feedback, I would have loved to have also a space for qualitative comment at the end of it. |
| | | | | | | I think it's because there is a personal bias going on. In fact, when you listen to a specific person, if you like her, or you have a relationship or something, your feedback is biased by them. And so in some cases, F and S were aligned, but maybe in another assessment F wasn't present, or did not have the same opinion and S or U. So it seems super interesting to discover how much we are biased by the fact that this tool provides objective answers. |
| | | | | | | I think that the first and the second time, I needed to, to, to arrange and find my positioning using it. So maybe the first. And the second time, I was a little bit biased by the fact that I didn't know exactly what to do that from the third or fourth, I was growing confident in using it. |
| | | | | 3.2.2. Definitely a barrier | *The person describes some feature of the tool as preventing from giving proper feeback or possibly invalidating it.* | It has sometimes been difficult to answer the questions, it was difficult to rate rate so quickly. For example, the engaging was difficult for me to read because I felt they had a difficult task and the task was to remember and not to really engage so. But there was a difference between the speeches, of course, the first speeches, it was more engaging. So I wasn't so sure about how to rate that because I thought the criteria were different. |
| | | | | | | Yeah, it was really quick. I could do it really quick. But then I was wondering, do I really do justice? |
| | | | | | | I agreed that I was a little bit uncomfortable at listening. And at the same time, going on another page and providing the feedback. You know, while I was providing the feedback, I did not have the time to listen to the speaker, because I was focused on the form and the not on the speech. |
| | | | | | | I also felt very uncomfortable. I first of all, I'm not used to giving like just a scoring in terms of number because I don't know if you get like you have a three out of five at the end. That's not really a feedback. It's just a number you don't really know How to improve or what to do differently at the end. Also, I think for me, like the first two or three were like trials, they were like setting the scene for me. And based on that, I gave scores to the next three ones. So there was also a little bit yeah, like, at the end, I was like, Okay, the first one, I should have scored differently. So I was starting to make some changes in my head. But at the same time, the form was already sent to you. |

Appendix 9: Main concepts of ESIT's Interpretive Theory



Appendix 10: Plan of the workshop

**Webinar plan**

| | | |
|---|---|---|
| **HOST:** Mikhail Demidov<br>**CO-HOST:** OAMK colleague | **DATE:** February 12, 2022 | **TOPIC/THEME:**<br>Giving structured feedback in student training |
| **EVENT:** OAMK | **TIME:** 15:30 – 16:30 Finnish time (60') | **WEBINAR FOCUS:**<br>Scaffolded assessment vs free assessment |

**AIMS** (main takeaways from the webinar)

**Primary:**

- For participants to get sensitized to the need for tools for giving structured feedback in student training.

**Secondary:**

- For participants to familiarize themselves with typical issues that students of interpreting have during the early phase of their training.

**Personal:**

- To present my project to a wider audience in a greater detail and, possibly, to expand my professional network;
- To set up and fit test some of the infrastructure (MS Forms, MS Power Automate, etc.) that may be required for practical part of my thesis.

**GROUP PROFILE**

**OAMK students:**

Post-degree learners with background in education and/or entrepreneurship. Many people are holding / have held positions in administration or management.

Several people are currently teaching. A few people are running their own business. The areas of expertise vary significantly from hospitality industry and language learning to project management and coding (programming).

In general, the group is very active and willingly engages in communication and practice-based exercises.

**ASSUMPTIONS**

- I assume the group will be largely unfamiliar with the issues of interpretation, however they will likely be interested in a short interactive activity;
- I assume the group will be familiar with various forms of competence / performance assessment, which should make our dialogue relevant;
- I assume I will need a colleague to co-host the event, as managing a one-off interactive activity online may put significant strain on a single host.

**SOURCES OF MATERIALS**

- **Turn-taking list**;
- Picture on how to wash hands (Handwashing instructions, CDC https://www.cdc.gov/handwashing/images/handwashing-day/2020/wash-your-hands-banner.jpg);
- **Poll 1** on MS Forms (assessment of Pre-task);
- **Recording 1** (1-minute speech for consecutive "interpretation" = 3 bits of 20 seconds each, with pauses) — the text should have clear logic but uncommon terminology, so as to allow the Speaker paraphrase the text in his/her own words. In general, the rendition of this text should still sound plausible, but retold by the speaker (https://webgate.ec.europa.eu/sr/speech/scientific-use-eggs);
- **Poll 2** on MS Forms (assessment of Rendition 1);
- **Recording 2** (1-minute speech for consecutive "interpretation" = 3 bits of 20 seconds each, with pauses) — the text should have clear words but less obvious logic (for example, the effects of cortisol on calcium levels — familiar words with unfamiliar links). The expected output is either a failed rendition or a very generalized rendition. Will likely produce the effect of an unplausible story (https://webgate.ec.europa.eu/sr/speech/interpreting-united-nations-office-geneva);
- **Poll 3** on MS Forms (assessment of Rendition 2);

**ROLES**

- **Speakers** (OAMK) —participants who agree (at the preliminary stage) to present a free speech, to describe a picture, or to repeat a text in English (provided by the host). Pretend to be interpreters.

- **Assessors** (OAMK) — participants who agree to actively assess the performance of speakers and are expected to make contributions during all of the discussions (unlike Observers).

  **Tool (+)** — aka *Experts T*, participants who will receive links to polls in MS Forms. The polls reflect the assessment categories of the Prototype. So these participants have the advantage of using a tool for structured feedback. Therefore, their contribution is expected to be *written*. They *should not* speak out during the discussions, as this will reveal the process of structured assessment and will create bias (participants who provide free and unstructured assessment will be biased toward making their assessments structured as well).

  **Tool (-)** — aka *Experts E*, participants who provide free and unstructured assessment. Their contribution is expected to be *oral*. They are expected to actively speak out during each round of discussions.

- **Observers** (OAMK, probably other participants of the event, if any) — participants who choose not to actively participate. They will view the event, and will speak out during the discussions in case they want.

| STAGE | STAGE AIM(S) | PROCEDURE | TIMING | TYPE OF INTERACTION / MATERIALS / SPEAKERS | ANTICIPATED PROBLEMS AND SOLUTIONS / THINGS TO KEEP IN MIND |
|---|---|---|---|---|---|
| Before the event | • To assign participants to various groups;<br>• To give individual instructions;<br>• To make sure the needed number of people will be | Based on their profiles and interests, learners from OAMK will be contacted.<br><br>They will be offered various roles during the meeting and, if they agree, will be assigned to various groups: Speakers, Tool (+) assessors, Tool (-) assessors.<br><br>Each group will be given individual instructions | at least 3 days before the meeting | Communication on Slack | |

| | | | | | |
|---|---|---|---|---|---|
| | available and make necessary arrangements if not. | | | | |
| Lead-in | • To establish interest in the topic;<br>• To establish context. | Thank for joining, remind why this activity takes place and ask concept checking questions (CCQs):<br><br>– Who is in group 1, 2, 3, etc.?<br>– What does group 1, 2, 3, etc. do?<br>– What are the tools that group 1, 2, 3, etc. uses?<br><br>(participants should know it from previous instructions) | 5'<br><br>15:30-15:35 | Communication on the event platform: Teams / Zoom / Howspace… | The event will take place toward the end of day 3 of F2F, so no specific or lengthy lead-in will be needed.<br><br>It is important to check the understanding of activities by each group, so concept checking questions should be asked. |
| Warm-up | • To break ice;<br>• To get going;<br>• To create a relaxed atmosphere of the activity. | To ask OAMK Speakers 1-3 to talk for a minute about a familiar and common subject, e.g. how to make a cup of tea with a tea bag. | ~3'<br><br>~15:35-15:38 | • Free speech (off the top of participant's head) – OAMK Speakers 1-3<br><br>• Other participants just observe | Instruct other participants just to observe, no assessment at this point. |
| Pre-translation | • To shift from warm-up to the core the activity. | To instruct participants to look at the picture and get ready to talk about it. (Ask CCQs). | ~7'<br><br>15:38-15:45 | • Co-host<br>• OAMK Speakers 2, 3 | • Don't forget to ask CCQs;<br>• Don't let the conversation (discussion) to go too lengthy at this point. |

| | | | | | |
|---|---|---|---|---|---|
| | | To share screen and show a picture on how to wash hands (Handwashing instructions) for about a minute.<br><br>To ask one OAMK Tool (-) Assessor to share their impressions of the two speeches. | | • OAMK Tool (-) Assessor | • To monitor turn-taking of Assessors using the Turn-taking list.<br><br>• Co-host keeps track of categories / wordings / aspects that Tool (-) Assessors use. This will be presented at the end to underline the commonalities / differences between structured and unstructured feedback. |
| Pre-translation | • To present Takeaway 1 | To instruct OAMK Tool (+) Assessors to fill in Poll 1 on MS Forms about the difference of the two speeches.<br><br>To give about one minute to do so.<br><br>To briefly review the results (maybe co-host could communicate with the participants in the meantime) and present Takeaway 1.<br><br>*<u>Takeaway 1</u> (add possible corrections based on the vote):*<br><br>*A simple text on familiar and common subjects sounds fluent, natural (= appropriate style), relaxed, consistent (=logical).* | 5’<br><br>15:45-15:50 | • Co-host<br>• Tool (+) Assessors | • If I look through the poll results and summarize them, I will inevitably be distracted from the activity. In order to keep pace, a co-host is expected to keep the participants engaged in the meantime (a relevant informal conversation).<br><br>• Should MS Power Automate fail (automatic filling-in of Prototype 1 based on MS Forms), the host (me) transfers data manually, which makes the role of the co-host even more important. |

| | | | | | |
|---|---|---|---|---|---|
| | | *A constraint (picture + less common subject) makes text sound less fluent, less natural (= appropriate style), less relaxed, but still consistent (= logical).* | | | |
| Translation (in the context of this event = repetition of text in English) | Rendition 1<br>• For speakers to get sensitized to issues of interpretation;<br>• For Tool (+) Assessors to try the tool for giving structured feedback;<br>• For Tool (-) Assessors to get the feel of unstructured feedback. | To instruct participants to listen to Recording 1 (played in 20-second bits; Speakers make their renditions during the pauses).<br><br>To instruct Speakers to get ready.<br>To instruct Tool (-) Assessors to get ready.<br><br>To provide Tool (+) Assessors with links to Poll 2. The poll can be completed during the rendition or immediately afterwards.<br><br>**Ask CCQs.**<br><br>To play the recording in bits. Speaker 2 is asked to make the rendition.<br><br>To give about a minute to complete Poll 2 [Tool (+) Assessors] or to prepare to speak out [Tool (-) Assessors].<br><br>To ask one OAMK Tool (-) Assessor to share their impressions of Rendition 1. | 10’<br><br>15:50 – 16:00 | • Recording 1<br>• OAMK Speaker 2<br>• OAMK Assessors | • Don’t forget to ask CCQs<br>• Should the playback of the recording fail for any reason, the co-host will be asked to read the speech out.<br>• Co-host keeps track of categories that Tool (-) Assessors use. [Probably at this stage the unstructured feedback will start to become vague, generally positive, and somewhat difficult to act upon — if this is the case, it is a very illustrative finding that should be presented at the end of the Workshop].<br>• To monitor turn-taking of Assessors using the Turn-taking list. |
| Translation (in the context of | To summarize the results of Rendition 1 | To briefly review the results (maybe co-host could communicate with the participants in the meantime) and present Takeaway 2. | 5’ | • Host<br>• Co-host | • In order to keep pace, a co-host is expected to keep the participants engaged in the |

| | | | | | |
|---|---|---|---|---|---|
| this event = repetition of text in English) | | *<u>Takeaway 2</u>: the text with clear logic but uncommon terminology allows the Speaker paraphrase the text in his/her own words. In general, the rendition of this text can still sound plausible, but produce the impression of being retold rather than rendered (or "translated" accurately) by the Speaker.* | 16:00-16:05 | | meantime (a relevant informal conversation) before the Takeaway is presented by the host. |
| Translation (in the context of this event = repetition of text in English) | Rendition 2<br>• For speakers to get sensitized to issues of interpretation;<br>• For Tool (+) Assessors to try the tool for giving structured feedback;<br>• For Tool (-) Assessors to get the feel of unstructured feedback. | To play the recording in bits. OAMK Speaker 3 is asked to make the rendition.<br><br>To give about a minute to complete Poll 3 [Tool (+) Assessors] or to prepare to speak out [Tool (-) Assessors].<br><br>To ask one OAMK Tool (-) Assessor to share their impressions of Rendition 2. | 10'<br><br>16:05 – 16:15 | • Recording 2<br>• OAMK Speaker 3<br>• OAMK Assessors | • Should the playback of the recording fail for any reason, the co-host will be asked to read the speech out.<br>• Co-host keeps track of categories that Tool (-) Assessors use.<br>[I expect the difference between structured and unstructured feedback to become even more evident by this point].<br>• Instruct everyone (if needed) *not* to go on with a more general discussion (as the next stage is allocated for it), but to stick to Rendition 2 only.<br><br>• To monitor turn-taking of Assessors using the Turn-taking list. |

| | | | | | |
|---|---|---|---|---|---|
| Translation (in the context of this event = repetition of text in English) | To summarize the results of Rendition 2 | To briefly review the results (maybe co-host could communicate with the participants in the meantime) and present Takeaway 3.<br><br>*<u>Takeaway 3</u>: the text with clear wordings but less obvious logic can lead to either a failed rendition or a very generalized rendition. This will likely produce the effect of an unplausible story. Some special interpreting techniques (beyond the scope of this workshop) are needed.* | 5’<br><br>16:15-16:20 | • Host<br>• Co-host | • In order to keep pace, a co-host is expected to keep the participants engaged in the meantime (a relevant informal conversation) before the Takeaway is presented by the host. |
| Post-translation | • For any participant to share any observations;<br>• For everyone to have a discussion;<br>• To summarize the results of the Workshop. | (1) To give the floor to the Speakers and ask about their experience;<br>(2) To give floor to Tool (+) assessors and ask about their experience;<br>(3) To give floor to observers and ask about their experience;<br>(4) To ask any other participants (if any) to share their observations;<br>(5) To ask co-host to share his/her observations;<br>(6) To share the results in the form of Prototype 1 (completed Excel report) and ask about the immediate feedback on this tool.<br>(7) To share Takeaway 4 and summarize the Workshop.<br>(8) Closing remarks.<br><br>*<u>Takeaway 4:</u> the process of consecutive interpretation encompasses:* | 10’<br><br>16:20-16:30 | Everyone | • To ask the co-host to take note of the insights that participants will share;<br>• To ask the co-host to present his/her observations [the results of ‘keeping track’ mentioned above];<br>• In the meantime, the host (me) prepares and shares the results of the Polls transferred to Prototype 1. |

| | | | | | |
|---|---|---|---|---|---|
| | | *(a) understanding; and*<br>*(b) rendering.*<br><br>*Rendering, in turn, encompasses two components:*<br>*(b1) memorizing or "internalizing" the text; and*<br>*(b2) reproducing the text.*<br><br>*Which is what Interpretive theory of Translation describes:*<br>*Understanding – De-verbalization —*<br>*Re-expression in a different language.* | | | |

Note: the required pedagogical materials are highlighted yellow; the activities by the co-host are highlighted purple.