



# MASTERARBEIT | MASTER'S THESIS

Titel | Title

The Relevance, Usability and Acceptance of Automatic Speech Recognition (ASR) in European Institutions: A Romanian-Language Case Study at the European Parliament

verfasst von | submitted by  
lic. Oana-Andreea Toader

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of  
Master of Arts (MA)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:

UA 070 348 331

Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:

Masterstudium Translation Italienisch Deutsch

Betreut von | Supervisor:

Univ.-Prof. Dragos Ioan Ciobanu PhD



## ABSTRACT (English)

As speech technologies move from the margins of innovation to the centre of multilingual professional communication, automatic speech recognition (ASR) has become more than a technical development: it has become an institutional and translational concern. This thesis examines the relevance, usability and acceptance of ASR in European institutional settings, with particular empirical attention to the European Parliament (EP), the Romanian language and the professional workflows of the Directorate-General for Translation (DG TRAD). It approaches ASR not merely as a speech-to-text system, but as a technology whose practical value is shaped by linguistic conditions, professional expectations and institutional constraints. Combining theoretical discussion with an exploratory case study, the thesis investigates three forms of ASR use: live transcription, asynchronous transcription and active ASR use, including voice-enabled interaction.

The empirical part draws on an interview with the Head of the EP Speech-to-Text Unit, an interview with an experienced proofreader from the Romanian Translation Unit of DG TRAD, a structured presentation for that unit, a follow-up try-out and a questionnaire involving its members. By combining expert insight, practitioner experience, direct experimentation and structured feedback, the study seeks to examine more concretely how ASR is developed, perceived and evaluated within European institutional contexts.

The findings suggest that the relevance, usability and acceptance of ASR vary across European institutional workflows, depending on function, risk level and the degree of human control. While the EP live Speech-to-Text platform is relevant as an accessibility service for immediate access, especially in less politically sensitive settings, Verba is more suitable for official transcription workflows within DG TRAD, where human revision before publication makes ASR output more acceptable. Although the integration of ASR into CAT-tool workflows remains limited, Romanian-language professionals in the European Parliament express a cautious but clear openness to future active use of ASR, provided that the tools are institutionally secure and approved.

Overall, the thesis argues that ASR in European institutions is no longer a distant possibility but an emerging reality. However, its institutional use has not yet fully matured or been integrated in all its forms, and its future in EU institutions depends on gradual, secure and workflow-sensitive integration.

## **ABSTRACT (Deutsch)**

Während Sprachtechnologien vom Randbereich technologischer Innovation zunehmend in das Zentrum mehrsprachiger professioneller Kommunikation rücken, ist die automatische Spracherkennung (ASR – Automatic Speech Recognition) nicht mehr nur eine technische Entwicklung, sondern auch ein institutionelles und translationsbezogenes Themenfeld. Die vorliegende Masterarbeit untersucht die Relevanz, Gebrauchstauglichkeit und Akzeptanz von ASR in europäischen institutionellen Kontexten, mit besonderem empirischem Fokus auf das Europäische Parlament (EP), die rumänische Sprache sowie die professionellen Arbeitsabläufe der Generaldirektion Übersetzung (GD TRAD). ASR wird dabei nicht nur als Speech-to-Text-System verstanden, sondern als eine Technologie, deren praktischer Nutzen durch sprachliche Bedingungen, professionelle Erwartungen und institutionelle Rahmenbedingungen geprägt ist. Durch die Verbindung einer theoretischen Auseinandersetzung mit einer explorativen Fallstudie untersucht die Arbeit drei Formen des ASR-Einsatzes: Echtzeittranskription, asynchrone Transkription sowie aktive ASR-Nutzung, einschließlich sprachgestützter Interaktion.

Der empirische Teil stützt sich auf ein Interview mit dem Leiter der Speech-to-Text Unit des EP, ein Interview mit einem erfahrenen Proofreader aus der rumänischen Sprachabteilung der GD TRAD, eine strukturierte Präsentation für diese Abteilung, eine anschließende Erprobung sowie einen Fragebogen unter deren Mitgliedern. Durch die Verbindung von Expertenwissen, berufspraktischer Erfahrung, direkter Erprobung und strukturiertem Feedback zielt die Studie darauf ab, konkreter zu erfassen, wie ASR in europäischen institutionellen Kontexten entwickelt, wahrgenommen und bewertet wird.

Die Ergebnisse deuten darauf hin, dass die Relevanz, Gebrauchstauglichkeit und Akzeptanz von ASR je nach institutionellem Arbeitsablauf in europäischen Institutionen variieren und von Funktion, Risikograd sowie dem Ausmaß menschlicher Kontrolle abhängen. Während die Live-Speech-to-Text-Plattform des EP als Barrierefreiheitsdienst für unmittelbaren Zugang relevant ist, insbesondere bei weniger politisch sensiblen Veranstaltungen, eignet sich Verba stärker für offizielle Transkriptionsabläufe innerhalb der GD TRAD, in denen eine menschliche Überarbeitung vor der Veröffentlichung die ASR-Ausgabe akzeptabler macht. Obwohl die Integration von ASR in CAT-Tool-gestützte Arbeitsabläufe begrenzt bleibt, zeigen die Mitglieder der rumänischen Sprachabteilung im EP eine vorsichtige, aber deutliche Offenheit gegenüber einer künftigen aktiven Nutzung von ASR, sofern die Tools institutionell sicher und zugelassen sind.

Insgesamt vertritt die Arbeit die Auffassung, dass ASR in europäischen Institutionen keine ferne Möglichkeit mehr, sondern eine entstehende Realität ist. Allerdings ist ihre institutionelle Nutzung noch nicht vollständig ausgereift oder in all ihren Formen integriert, und ihre Zukunft in den EU-Institutionen hängt von einer schrittweisen, sicheren und an konkrete Arbeitsabläufe angepassten Integration ab.

# Table of Contents

---

|                                                                                                            |           |
|------------------------------------------------------------------------------------------------------------|-----------|
| <b>List of Figures and Tables.....</b>                                                                     | <b>7</b>  |
| <b>List of Abbreviations.....</b>                                                                          | <b>8</b>  |
| <b>1. Introduction .....</b>                                                                               | <b>9</b>  |
| <b>2. Theoretical foundation .....</b>                                                                     | <b>10</b> |
| 2.1 Automatic speech recognition (ASR): overview.....                                                      | 10        |
| 2.2 ASR in the professional translation process.....                                                       | 14        |
| 2.3 ASR in the European institutions .....                                                                 | 16        |
| 2.4 ASR and the Romanian language community .....                                                          | 23        |
| 2.5 Future directions in ASR: rationale and research gaps.....                                             | 26        |
| <b>3. Research questions and methodology .....</b>                                                         | <b>28</b> |
| 3.1 Interview with the Head of the European Parliament Speech-to-Text Unit .....                           | 30        |
| 3.2 Interview on Verba with an experienced user from the European Parliament Romanian Unit (DG TRAD) ..... | 32        |
| 3.3 Presentation for the European Parliament Romanian Unit (DG TRAD).....                                  | 33        |
| 3.4 Follow-up try-out (EP RO Unit, DG TRAD).....                                                           | 38        |
| 3.5 Questionnaire (EP RO Unit, DG TRAD) .....                                                              | 38        |
| 3.6 Methodology summary .....                                                                              | 39        |
| <b>4. Findings.....</b>                                                                                    | <b>40</b> |
| 4.1 Findings from the interview with the Head of the EP STT Unit.....                                      | 40        |
| 4.2 Findings from the Verba interview (EP RO Unit, DG TRAD).....                                           | 47        |
| 4.3 Findings from the presentation for the EP RO Unit, DG TRAD .....                                       | 53        |
| 4.4 Short findings from the follow-up try-out.....                                                         | 57        |
| 4.5 Findings from the questionnaire (EP RO Unit, DG TRAD).....                                             | 58        |
| <b>5. Concluding observations .....</b>                                                                    | <b>65</b> |
| 5.1 Summary of key findings in relation to the literature.....                                             | 66        |
| 5.2 Strengths and limitations.....                                                                         | 73        |
| 5.3 Practical implications and recommendations .....                                                       | 76        |
| 5.4 Final conclusion.....                                                                                  | 80        |
| <b>6. Bibliography .....</b>                                                                               | <b>82</b> |
| <b>7. Appendices .....</b>                                                                                 | <b>86</b> |

|                                                                        |            |
|------------------------------------------------------------------------|------------|
| <b>Appendix A - Interview with the Head of EP STT Unit.....</b>        | <b>86</b>  |
| Interview questions.....                                               | 86         |
| Transcript and recordings .....                                        | 87         |
| <b>Appendix B - Interview on Verba (EP RO Unit, DG TRAD) .....</b>     | <b>97</b>  |
| Interview questions.....                                               | 97         |
| Transcript and recording.....                                          | 97         |
| <b>Appendix C - Presentation for the EP RO Unit, DG TRAD .....</b>     | <b>102</b> |
| Presentation video .....                                               | 102        |
| Texts used .....                                                       | 102        |
| Explanation transcript.....                                            | 103        |
| <b>Appendix D - Follow-up questionnaire (EP RO Unit, DG TRAD).....</b> | <b>109</b> |
| Questionnaire .....                                                    | 109        |
| Questionnaire results .....                                            | 109        |
| <b>Acknowledgements.....</b>                                           | <b>110</b> |

## List of Figures

---

|                                                                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1: System notification on an EP laptop restricting Windows voice typing .....                                                                        | 19 |
| Figure 2: Interface examples of the EP’s Speech-to-Text platform (Mavrič 2022: 35).....                                                                     | 21 |
| Figure 3: Disclaimer preface and language-selection screen of the EP’s Speech-to-Text platform (Mavrič 2022: 36).....                                       | 21 |
| Figure 4: Screenshot from the presentation video (Appendix C) showing an email drafted in Romanian with Windows 11 voice typing .....                       | 34 |
| Figure 5: Screenshot from the presentation video (Appendix C) showing Dragon Professional within Trados Studio desktop.....                                 | 36 |
| Figure 6: Screenshot from the presentation video (Appendix C) showing Trados Online Editor with Romanian dictation and voice commands.....                  | 37 |
| Figure 7: Excerpt from the verbatim report of proceedings, 9 February 2026, Strasbourg (European Parliament 2026a).....                                     | 48 |
| Figure 8: Video-on-demand entry for Vlad VASILE-VOICULESCU, Sitting of 9 February 2026 (European Parliament 2026b).....                                     | 49 |
| Figure 9: Email from DG TRAD Service Desk confirming Windows voice typing restrictions (dgtrad.servicedesk@europarl.europa.eu).....                         | 54 |
| Figure 10: Email from DG TRAD Service Desk confirming different rules for short-term and long-term Dragon use (dgtrad.servicedesk@europarl.europa.eu) ..... | 55 |

## List of Tables

---

|                                                                                                          |    |
|----------------------------------------------------------------------------------------------------------|----|
| Table 1: Transcribed speech hours across languages in the VoxPopuli corpus (Wang et al. 2021: 994) ..... | 17 |
| Table 2: Overview of the availability of Romanian speech corpora (Georgescu et al. 2020: 6607) .....     | 23 |

## List of Abbreviations

|         |                                                             |
|---------|-------------------------------------------------------------|
| AI      | Artificial intelligence                                     |
| ASR     | Automatic speech recognition                                |
| CAT     | Computer-assisted translation                               |
| DG TRAD | Directorate-General for Translation                         |
| DTW     | Dynamic time warping                                        |
| ENVI    | Committee on the Environment, Public Health and Food Safety |
| EP      | European Parliament                                         |
| EU      | European Union                                              |
| LDS     | European Language Data Space                                |
| MEP     | Member of the European Parliament                           |
| PV-CRE  | Procès-verbaux et comptes rendus de la plénière             |
| STT     | Speech-to-Text                                              |
| VRCP    | Voice Recognition Call Processing                           |



# 1. Introduction

In a language industry increasingly shaped by speech technologies embedded in phones, cars, software, web applications and customer-service systems, automatic speech recognition (ASR) has become a visible presence in everyday life, professional language work and Translation Studies research. Building on more than five decades of technological development, from early speech-to-text systems to more complex forms of speech-enabled human-machine interaction, ASR now raises important questions not only about technical performance, but also about practical usefulness in specific professional contexts.

This thesis investigates ASR in a more targeted way by examining its relevance, usability and acceptance in European institutions through a Romanian-language case study at the European Parliament (EP). In this context, ASR is approached as part of broader forms of speech-enabled interaction in professional environments, whose value depends on how effectively they respond to real workflow needs, language-specific conditions and institutional requirements.

The thesis is situated within the broader context of ASR use in European institutions, while its empirical focus lies specifically on the European Parliament and its related workflows. This setting brings together several distinct forms of ASR use within the same multilingual institution, including real-time transcription for accessibility, internal transcription for documentation purposes and active dictation or voice-enabled interaction in everyday office and translation-related tasks. It therefore offers a useful case for examining ASR not only as a technical system, but also as a workflow-dependent technology shaped by institutional rules, professional expectations and language-specific conditions.

More specifically, the thesis asks how current institutional ASR solutions in the European Parliament are described and evaluated by experienced professionals, how ASR may be integrated into translation and CAT-tool workflows under organisational constraints and how Romanian-language professionals perceive ASR in practice.

To address these questions, the thesis first outlines the theoretical foundations necessary for understanding ASR as a technological, translational and institutional phenomenon. It then combines this broader contextual discussion with an empirical investigation of three forms of ASR use within the EP: live transcription, asynchronous transcription and active dictation or voice-enabled interaction. This investigation draws on interviews, a structured presentation, a follow-up try-out and a questionnaire with Romanian-language professionals from the Directorate-General for Translation (DG TRAD).

In this way, the thesis seeks to contribute to a better understanding of the current role of ASR in different institutional workflows and to identify possible directions for future integration strategies. More specifically, it adopts an exploratory institutional case-study perspective in which ASR performance matters insofar as it affects workflow fit, correction effort, institutional feasibility and user acceptance.

## 2. Theoretical foundation

Understanding the role of automatic speech recognition (ASR) in professional and institutional workflows requires a broader perspective on the technology's development, its integration into translation-related practice and its current framing within European institutions. Against this background, the Romanian context can be examined more clearly, together with the research gaps that motivate the present study.

### 2.1 Automatic speech recognition (ASR): overview

ASR is now an everyday presence, often operating in the background of private and professional devices. Users dictate messages while driving, ask an iOS or Android voice assistant to perform various tasks on their phone, or speak on a meeting platform that appears to produce transcripts effortlessly. However, behind this apparent simplicity and these everyday uses lies a technically demanding task. Over the last five decades, this technology has developed with the goal of “converting a speech waveform (as an acoustic realization of a linguistic event) into words (as a best-decoded sequence of linguistic units)”, thereby enabling systems not only to produce text output but also to support actions (Juang & Rabiner 2004: 18).

This broad definition suggests that: what counts as ASR has expanded as expectations and the settings in which it is used have changed, with direct implications for perceived relevance and usability. Juang and Rabiner (2004) note that early ASR research focused first on recognising isolated spoken words, then on speech-to-text systems, and later on treating this process as a necessary first step toward machines that can understand and respond to human speech. In other words, the concept of ASR has evolved together with its use cases, shaping decades of research and shifting progress.

Early work in the 1950s treated ASR primarily as proof of feasibility under strictly controlled conditions. In his survey, Furui reports that, in 1952 at Bell Laboratories, Davis, Biddulph and Balashek built a system for isolated digit recognition for a single speaker, using formant frequencies measured or estimated during the vowel regions of each digit (Furui 2005: 64–65). This early success is often cited because it marked an important starting point: it demonstrated that speech recognition was possible, but only for isolated words, a single speaker and a limited vocabulary.

As research moved beyond isolated digits to connected words, ASR became a problem of variability management. The same speaker can produce the same words at different speeds, with different pauses or articulations. Without technical adjustment, the acoustic patterns may not align well enough for the system to recognise them as the same word. Furui highlights this “non-uniformity of time scales” in speech events and points to dynamic time warping (DTW) as a key response: a time-normalisation method that aligns two speech patterns despite differences in speaking rate. Juang and Rabiner likewise describe DTW as a formal pattern-matching technique that became essential for later developments. From a usability perspective, DTW matters because it reflects a design logic that still shapes modern ASR: systems must

adapt to human speech variability rather than require users to speak in an unnaturally standardised way (Furui 2005; Juang & Rabiner 2004).

This awareness of variability enabled the transition from connected words to continuous speech. In connected-word recognition, systems typically assume short sequences of words with relatively clear boundaries, often signalled by pauses, which makes segmentation and matching easier. In continuous speech, words run together. Boundaries are not marked, co-articulation blurs adjacent sounds and speech rate fluctuates. Continuous speech recognition is qualitatively harder because the system must effectively decide where one word ends and the next begins while decoding the signal. It became possible once ASR stopped relying mainly on matching whole spoken words to stored templates and instead searched for the most likely word sequence directly from the audio signal. In practice, the system evaluates many possible word strings and uses language constraints - that is, likely word sequences - together with the acoustic evidence to decide both which words were said and where one word ends and the next begins, all at the same time (Arora & Singh 2012).

In realistic speech conditions, clear boundaries do not exist. Pauses, rephrasing, changes in speaking speed and incomplete phrases are normal. This leads to a key question: how well does today's ASR handle this kind of disfluent speech in real professional settings? A translator may dictate while thinking aloud and reformulating. A speaker in a parliamentary debate may produce fast, interrupted or unclear speech. In such cases, the issue is whether the result is usable (with a reasonable correction effort), reliable (with errors that are predictable and manageable) and therefore acceptable for professional work.

Another need for speech recognition under realistic conditions drove the shift from speaker-dependent systems to speaker-independent use. Early successes focused on trained speakers, but deployment in professional and private settings required systems that work across voices, accents and speaking styles without requiring every user to train them. Juang and Rabiner make this motivation concrete in telecommunication settings: ASR was expected to serve a very large user population without individual training, which pushed research toward generalisation across speakers. To serve many voices, researchers began collecting speech from many talkers and then used speech clustering algorithms to build reference patterns that represent groups of similar pronunciations, first as multiple templates and later as statistical prototypes (Juang & Rabiner 2004: 11).

From an acceptance perspective, this is an important historical development. If effective use depends on extensive personal training, adoption becomes difficult in time-pressured, multilingual institutional workflows. At the same time, the shift to speaker-independent ASR raises a reverse question: even when systems no longer require training the machine to a user's voice, do users still end up training themselves to the system?

In other words, professional uptake may depend on whether users feel they must modify their pronunciation or articulation to reduce recognition errors and minimise correction effort, effectively adapting their speech behaviour to improve the written output. This can slow down drafting, interrupt cognitive flow and reduce perceived usefulness, even when the technology is no longer speaker dependent.

In parallel with speaker-independent ASR, the technology evolved from small vocabularies to large-vocabulary dictation and documentation. With large vocabularies, many word sequences become plausible and the system must rely more heavily on linguistic constraints to guide decoding. Juang and Rabiner discuss this logic in the context of dictation-oriented development and associate this progress with the use of statistical language modelling, such as n-grams, to predict likely word sequences. They explain that a language model can be represented as a set of statistical rules describing how likely a given sequence of language symbols is. An n-gram model is a probabilistic mechanism for predicting word sequences (Juang & Rabiner 2004: 10). However, a large vocabulary is not the same as a domain-appropriate vocabulary. Even if acoustic decoding improves, domain mismatch can still produce systematic errors when institutional phraseology, proper names, acronyms and specialised terminology diverge from general-language expectations. Usefulness therefore depends on whether the language model can be adapted to an institutional context. For example, can an ASR system be trained to distinguish between ENVI (the European Parliament's Committee on the Environment, Public Health and Food Safety) and the ordinary word *envy*?

Together with these developments, the meaning of ASR also shifted from speech-to-text for transcription purposes to voice-driven systems that can respond or act. Juang and Rabiner show that the telecommunication settings mentioned above also pushed ASR in this direction. At AT&T Bell Labs, ASR was developed for automated public services such as voice dialling and command-and-control for call routing, where systems were expected to work for millions of talkers without individual speaker training. A key concept was keyword spotting, that is, detecting a key word or phrase embedded in a longer natural sentence and using it to trigger a predefined action such as routing. A concrete illustration is the AT&T Voice Recognition Call Processing (VRCP) service introduced in 1992, which used ASR to route and handle calls at very large scale (Juang & Rabiner 2004: 3–12).

This development also changes what “good performance” means. In dictation, ASR is mainly judged by the quality of the text it produces and by how much correction is needed. In voice commands, what matters most is whether the system understands what the user wants and carries out the right action. In a CAT-tool environment, this distinction becomes concrete because ASR can serve two workflow functions at once: dictation (speech-to-text) to produce a draft translation for the active segment, and command-and-control (speech-to-action) to operate the interface by voice. The relevant question is not only whether ASR recognises spoken input accurately enough to dictate a translation, but also whether it reliably recognises voice commands that allow translators to navigate segments, confirm actions or insert standard operations more efficiently than with the mouse and keyboard.

In their book, Yu and Deng argue that speech became a common way to interact with technology only after ASR systems became usable enough for real-life situations and once many devices made keyboard and mouse input less convenient. They link this change to greater computing power, much larger amounts of real-world data, and the spread of mobile and embedded devices, where speech can be faster and easier. They also suggest a helpful way to group applications: ASR supports both human-human communication (for example, transcription and messaging) and human-machine communication (for example, voice search, digital assistants and in-vehicle systems) (Yu & Deng 2015: 1–3). For this thesis, this classification helps clarify that ASR used for dictation (drafting a text), transcription

(documentation) and voice control (interface navigation) represent different workflow functions. These functions may involve different acceptance dynamics among professional users and across institutional workflows.

Further, artificial intelligence has brought ASR to a more advanced stage of development. In this context, AI refers to the broader use of data-driven computational methods, while deep learning is one of the main AI approaches behind recent progress in speech recognition. Kheddar, Hemis and Himeur explain that recent progress in ASR is closely linked to methods that allow systems to learn more effectively from speech data and to perform better under difficult real-world conditions. In particular, they note that deep transfer learning helps build high-performing ASR models even when only small but related datasets are available; federated learning makes it possible to train models on sensitive or confidential speech data without centrally collecting the data themselves; reinforcement learning supports better decision-making in dynamic environments and can reduce computational costs; and transformers are valuable because they can capture broader dependencies within speech sequences (Kheddar *et al.* 2024). In this sense, artificial intelligence has improved ASR by making the technology more adaptable, more privacy-aware and better suited to multilingual, real-world use.

Driven by deep-learning-based AI methods and the availability of large amounts of real-world data, ASR has become an everyday technology. This shift from early laboratory systems to today's daily use also helps explain why the literature uses several overlapping terms for similar practices.

When the goal is to produce written text, authors use terms such as speech-to-text, voice typing, dictation or automatic transcription. When the goal is to trigger actions or support interaction, the same recognition component is discussed as voice commands, voice assistants or voice search. Because this thesis considers both the conversion of speech into text and voice-based interaction, it adopts automatic speech recognition (ASR) as an umbrella term for converting speech into machine-usable linguistic representations, while more specific labels (dictation, transcription, voice commands, etc.) are used whenever it is important to distinguish the workflow function and its practical implications.

This overview of ASR development has traced how the technology moved from early, controlled demonstrations to systems designed for real-world use: from isolated and connected words to continuous speech, from speaker-dependent setups to speaker-independent operation, from small vocabularies to large-vocabulary dictation and interactive services, and finally to AI- and deep-learning-based systems better suited to multilingual and real-world conditions.

In other words, the history of Automatic Speech Recognition (ASR) is a story of increasing technical coverage. However, technical progress that may be useful in theory does not automatically translate into professional uptake.

## 2.2 ASR in the professional translation process

In the professional translation process, the idea that technical progress in ASR does not automatically lead to professional uptake is especially relevant, because translation involves more than simply dictating a target text. It is a constrained activity shaped by deadlines, quality standards, CAT-tool interfaces and domain-specific language. In this sense, translation workflows place ASR under particularly demanding conditions. These conditions make visible the difference between ASR that works in principle and ASR that is perceived as relevant, usable and trustworthy in a particular professional practice.

Zapata *et al.* argue that speech technologies have reached a stage of development at which they are usable in several domains and note that interest also in translation-related ASR applications is increasing among researchers, trainers and professional practitioners. However, they stress that the translation-technology market has been slow to implement speech technologies in common CAT environments, which helps explain why ASR remains only partly integrated into everyday translation practice (Zapata *et al.* 2023: 16–17). Building on this observation, the present thesis aims to update the picture by examining the state of ASR integration in CAT tools at the time of writing (2025–2026). The focus is on the ASR functionality available for Romanian and its practical use in Trados Studio, a widely used CAT tool within European institutional translation workflows.

Even when ASR is used, it is rarely used on its own. Ciobanu and Secară make further clear that professional use is rarely only dictation. It is typically hybrid, meaning ASR is combined with keyboard and mouse and is used only in certain phases of the workflow. In their findings, only one respondent reported operating the computer solely by speech, while most used ASR mainly to dictate a draft and then switched back to keyboard and mouse for editing (Ciobanu & Secară 2019: 97).

For the present thesis, an important point is that this hybrid pattern is often described in only one direction: speak a draft, then edit it by typing. In institutional translation, however, translators often start from pre-existing text, such as translation-memory or machine-translation suggestions, and spend a large part of their work revising. This raises the need for more research on how speech technologies interact with revision and review practices. ASR should therefore be assessed not only as a drafting tool, but also as a possible support for revision work, including through voice-assisted corrections and navigation.

Another important issue is efficiency. A second insight from Ciobanu and Secară is that users often cite speed as a motivation and that some report substantial productivity gains (approximately 35% in the findings discussed), but they also outline that gains depend on recognition quality: if accuracy is too low, correction time can cancel out the benefit (Ciobanu & Secară 2019: 96–97). Other experimental work regarding the dictation process supports this statement. In their research, Zapata and Søbørg Kirkedal conclude that ASR accuracy must be above around 83.3% for dictation to become more efficient than typing and they show that small differences in accuracy can lead to large differences in efficiency (Zapata & Søbørg Kirkedal 2015: 207–208).

However, much of the discussion and many experimental results are based on a small number of languages and on controlled settings, so the findings cannot automatically be applied

to all professional contexts. Building on Zapata and Søeborg Kirkedal's point that dictation becomes more efficient than typing only when accuracy is high enough, this thesis examines how much correction effort is needed for Romanian in a European institutional environment. In this kind of setting, dense terminology, proper names and formulaic phrasing can lead to recurring, domain-related ASR errors rather than occasional isolated mistakes. Overall, the existing overview literature does not fully offer a Romanian-focused, institution-based comparison between what is technically available and what is actually relevant, usable and adopted in practice.

Ergonomics adds another layer to acceptance. Ciobanu and Secară also associate dictation with a potentially more comfortable physical working setup. Depending on the setup, translators may keep their eyes on the source text for longer and maintain a more relaxed posture, and some may even move more freely while translating with ASR. However, they also show that most translators still combine ASR with the keyboard and mouse and often use ASR mainly for the first draft, switching back to traditional input for editing. This suggests that ergonomic benefits are possible, but they do not automatically lead to a fully hands-free workflow (Ciobanu & Secară 2019: 97–101). In an institutional context, professional translators typically work in an office setting, alone or with only a small number of colleagues. At the time of writing, European Parliament translators in Luxembourg also have several days of teleworking per week. Under these conditions, dictation may be more feasible than in open-plan environments, because speaking to the computer is less likely to disturb others and can be better controlled, for example by closing a door or working from home.

In institutional settings, the use of voice input is shaped not only by usability but also by practical constraints, especially whether speaking aloud is appropriate in the workplace and compatible with confidentiality requirements. In the European Parliament, ASR is not just a matter of personal preference. Staff are bound by a “duty of discretion and must refrain from any unauthorised disclosure of information obtained in the line of duty” (European Parliament 2008: 6). This means that dictation can be problematic when translators work in shared or easily overheard environments, for example with an open door, while handling confidential documents. It also raises an additional issue beyond speaking aloud: where speech data are processed and stored and who may have access to them.

These requirements interact directly with the technical setup of many ASR solutions, particularly when cloud-based processing and storage are involved in research or commercial tools (Zapata *et al.* 2017: 116–117). According to the Annual Activity Report of the European Parliament's Directorate-General for Translation (2024: 31–43), data-protection compliance in DG TRAD is ensured through documented data-processing records, while IT governance places strong emphasis on access management and cybersecurity controls. Although these constraints are often mentioned in general terms, there is still limited empirical evidence on which ASR solutions used in translation workflows can, or could, comply with such European institutional requirements in practice.

Overall, the literature suggests that ASR is technically mature and increasingly discussed in translation, but it is still only partly integrated into CAT workflows. Where it is used, it is typically hybrid and combined with keyboard and mouse input. Its efficiency depends on recognition quality and on the correction effort required.

However, existing evidence remains limited for Romanian and for European institutional conditions, where dense terminology can lead to recurring domain-specific errors and where the practical use of ASR is further shaped by restrictions related to confidentiality, data protection, IT governance and tool availability on institutional devices. This subchapter therefore motivates examining, under 2025–2026 conditions, which ASR options available in a CAT-tool workflow can realistically meet these institutional constraints and how professional users perceive their usability and acceptability under these rules.

## 2.3 ASR in the European institutions

While translators may adopt ASR only occasionally in their individual workflows, European institutions increasingly use ASR as infrastructure for multilingual communication, accessibility and time-critical language mediation. This shift of perspective, from ASR as an optional input method at the translator’s desk to ASR as an institutional service, changes both the success criteria and the constraints. The relevant question is whether ASR can operate reliably, securely and consistently across languages, services and users under real institutional conditions.

Building on the distinction introduced in 2.1, ASR as support for human–human communication (capturing, indexing, sharing speech) versus human–machine interaction (speech as an input channel for operating tools), this thesis uses a more detailed categorization. In order to describe European institutional uses of automatic speech recognition, it will be distinguished between: ASR in public context (captioning, accessibility, streaming support), internal ASR services (transcription for workflows and documentation) and ASR capacity building (projects that create training data and reusable infrastructure). This helps explain an apparent distinction in terms of relevance, usability and acceptance. ASR may, for example, function successfully as an institutional service, while remaining only rarely used at the translator’s workstation.

### *European Commission: from capacity building to internal ASR services*

The European Commission presents ASR as part of Europe’s “critical digital capacities”. In a Digital Europe Programme project description on its website, the Commission outlines three main actions: a study of existing ASR solutions and the availability of training data, an open-source client–server streaming ASR prototypes for SMEs and public administrations, and the collection of datasets for selected low-resourced EU languages, namely Czech, Estonian and Greek (European Commission 2024)<sup>1</sup>. This focus on building internal capacity also highlights an important requirement for EU institutions: ASR solutions must meet internal governance rules about how speech data is processed, where it is stored and who can access it. In practice, institutions need control over the data, reproducible results and clear compliance, particularly in these contexts where confidentiality matters and where institutions aim to avoid sending sensitive content to commercial cloud services with unclear processing conditions.

---

<sup>1</sup> <https://digital-strategy.ec.europa.eu/en/news/automated-speech-recognition>



Furthermore, the Commission also outlines on its page that progress in ASR depends on “the availability of training data (speech and related transcription)” (European Commission 2024). For EU institutions, the key resource is in-domain institutional speech with aligned transcripts. This is what enables reliable training, evaluation and domain adaptation. As a result, multilingual ASR capacity is likely to develop unevenly across languages, not necessarily because of available ASR data in general, but because transcribed institutional speech is not equally available for all languages. Evidence of this domain imbalance is visible in VoxPopuli, one of the best-known corpora based on European Parliament recordings. The corpus draws on EP event recordings from 2009 to 2020, including plenary sessions, committee meetings and other events, and shows that the transcribed component is much larger for some languages, for example English (543 hours), German (282 hours), French (211 hours) and Spanish (166 hours), while others have substantially less coverage. Romanian has 89 hours of transcribed speech. The authors also note that speeches are only partially transcribed and that the raw data include missing audio, incomplete transcripts and inaccurate timestamps, all of which require filtering and processing (Wang et al. 2021: 994–995, Table 1).

|     | Unlab.<br>Hrs | Hrs  | Transcribed<br>Spkrs (F%) | Tkns | LM<br>Tkns |
|-----|---------------|------|---------------------------|------|------------|
| En  | 24.1K         | 543  | 1313 (29.6)               | 4.8M | 60.1M      |
| De  | 23.2K         | 282  | 531 (30.6)                | 2.3M | 50.0M      |
| Fr  | 22.8K         | 211  | 534 (38.6)                | 2.1M | 58.6M      |
| Es  | 21.4K         | 166  | 305 (40.6)                | 1.6M | 57.4M      |
| Pl  | 21.2K         | 111  | 282 (23.7)                | 802K | 13.6M      |
| It  | 21.9K         | 91   | 306 (33.8)                | 757K | 52.1M      |
| Ro  | 17.9K         | 89   | 164 (27.6)                | 739K | 10.3M      |
| Hu  | 17.7K         | 63   | 143 (30.3)                | 431K | 13.0M      |
| Cs  | 18.7K         | 62   | 138 (24.9)                | 461K | 13.5M      |
| Nl  | 19K           | 53   | 221 (39.3)                | 488K | 54.6M      |
| Fi  | 14.2K         | 27   | 84 (56.8)                 | 160K | 34.5M      |
| Hr  | 8.1K          | 43   | 83 (33.1)                 | 337K | 285K       |
| Sk  | 12.1K         | 35   | 96 (33.8)                 | 270K | 13.3M      |
| Sl  | 11.3K         | 10   | 45 (43.9)                 | 76K  | 12.6M      |
| Et  | 10.6K         | 3    | 29 (43.7)                 | 18K  | 11.3M      |
| Lt  | 14.4K         | 2    | 21 (14.8)                 | 10K  | 11.5M      |
| Pt  | 17.5K         | -    | -                         | -    | -          |
| Bg  | 17.6K         | -    | -                         | -    | -          |
| El  | 17.7K         | -    | -                         | -    | -          |
| Lv  | 13.1K         | -    | -                         | -    | -          |
| Mt  | 9.1K          | -    | -                         | -    | -          |
| Sv  | 16.3K         | -    | -                         | -    | -          |
| Da  | 13.6K         | -    | -                         | -    | -          |
| All | 384K          | 1791 | 4295                      | 15M  | 467M       |

Table 1: Transcribed speech hours across languages in the VoxPopuli corpus (Wang et al. 2021: 994)

This supports a key point: Romanian may not be low-resource in general, yet it can still be lower-resource in-domain in an institutional sense, where high-quality ASR requires strong coverage of specific names, acronyms, formulaic phrasing and specialised terminology. Even though Romanian is not among the languages explicitly targeted as “low-resourced” in the Commission initiative mentioned above, this thesis aims to clarify this point by investigating, through its research methodology, how Romanian is actually supported in current ASR capacity-building efforts within EU institutions.

The Commission's capacity-building work is complemented by internal, user-facing services. On its page on AI translation and language tools, the Commission lists, among others, a Speech-to-Text tool that allows users to upload audio or video files for full transcriptions or subtitles and is available in all EU official languages and some others. Access requires an EU Login account and registration, and the Commission describes the tools as free and secure AI language tools for eligible users (European Commission n.d.a)<sup>2</sup>.

It also states that user data are handled in line with EU data-protection rules and are not used to train commercial AI models. This controlled-access setting is especially relevant for institutional contexts, because the service is framed as a trusted EU-governed environment rather than as an unregulated public web service. The Commission therefore largely presents ASR at the level of infrastructure and internal service rather than at the level of translator workstations (European Commission n.d.b)<sup>3</sup>.

However, if institutions invest in ASR capacity-building and provide internal tools, it becomes important to examine to what extent these developments lead to routine ASR use in CAT environments at workstation level, if at all.

### ***Institutional constraints: confidentiality, data protection and access control***

Another important difference between institutional and general-market ASR is that relevance, usability and acceptance are inseparable from compliance. As mentioned in subchapter 2.2, officials and other servants are bound by a duty of discretion, meaning that dictating the content of sensitive documents is risky, especially when speech may be processed externally or stored in unclear ways. In this context, the question is whether ASR can be used within institutional rules on confidentiality, data protection and IT security.

DG TRAD's governance reporting makes this internal operational dimension visible. According to its Annual Activity Report, data-protection compliance is handled through concrete measures such as maintaining and validating data-processing records and providing staff support on data-protection questions. The same report also describes concrete information-security measures, including access management and restrictions on access to key systems (European Parliament, Directorate-General for Translation 2024: 42–43).

A concrete example is Windows 10/11 voice typing, a built-in speech-to-text function that can normally be used in many desktop applications for tasks such as drafting emails or taking notes. According to Microsoft Windows support documentation, the feature is available in Windows 10 and 11 and can be activated in supported environments (Microsoft Windows n.d.). At the time of writing this thesis, however, this function was restricted on EP-managed laptops. When attempting to activate Windows voice typing on an EP-managed Windows laptop, a system notification appears, stating that the speech service is managed by the organisation and that the feature is not available (Figure 1).

---

<sup>2</sup> <https://language-tools.ec.europa.eu/>

<sup>3</sup> [https://translation.ec.europa.eu/language-data-and-ai-using-ai-break-down-language-barriers/ai-based-multilingual-services-using-eu-language-data-innovate\\_en](https://translation.ec.europa.eu/language-data-and-ai-using-ai-break-down-language-barriers/ai-based-multilingual-services-using-eu-language-data-innovate_en)



*Figure 1: System notification on an EP laptop restricting Windows voice typing*

From an acceptance perspective, these confidentiality and security measures shape which ASR solutions can realistically be deployed and routinely adopted in the workplace. If a tool is unavailable on institutional devices because it depends on external cloud processing, it cannot easily become part of everyday professional practice, even if it could perform well in theory. This also helps explain why the distinction between general-market ASR and institutional ASR is central to the present thesis. In the EP context, users cannot simply choose any widely available dictation solution. They work within a controlled technical environment in which some tools are restricted.

### ***European Language Data Space: exchange of best practices***

A second signal comes from the European Language Data Space (LDS), a Commission-supported platform intended to enable the secure exchange and reuse of language data and related resources under EU governance and compliance conditions (LDS n.d.)<sup>4</sup>. It hosted a workshop dedicated to the “challenges and opportunities in Automatic Speech Recognition (ASR) solutions”, with the aim of exchanging best practices across Member States and fostering collaboration. The programme also indicates which issues institutions treat as central for institutional ASR: panels on how to obtain speech data, through crowdsourcing and cooperation with broadcasters, as well as panels on “Leveraging Parliamentary Speech Data” and “Addressing Legal Aspects” (LDS 2025)<sup>5</sup>.

On the one hand, parliamentary speech is treated as a major source of multilingual institutional audio, but it is also a challenging domain because it typically involves many speakers, formal registers, dense terminology and frequent proper names, features that can lead to systematic, domain-related errors rather than isolated mistakes. On the other hand, the inclusion of a panel on legal aspects highlights that reuse and deployment are shaped by

<sup>4</sup> [https://language-data-space.ec.europa.eu/index\\_en](https://language-data-space.ec.europa.eu/index_en)

<sup>5</sup> [https://language-data-space.ec.europa.eu/events/lds-workshop-speech-recognition-solutions-2025-07-01\\_en](https://language-data-space.ec.europa.eu/events/lds-workshop-speech-recognition-solutions-2025-07-01_en)

governance requirements. In practice, legal aspects commonly include questions such as usage rights, conditions for data sharing and rules for handling personal data and responsibility.

Overall, the LDS workshop reinforces key points already introduced by the Commission: how speech data are obtained and governed, which data are selected, how they are prepared and how services are integrated. The emphasis on sharing best practices further suggests that institutional success depends on how institutions collaborate and share solutions safely within controlled environments. At the same time, the workshop programme does not clarify to what extent these institutional efforts rely on in-house solutions or on partnerships with commercial providers in order to strengthen the development of ASR tools in EU institutions.

### ***The European Parliament: ASR as an internal service and public context tool***

Against this broader EU-level background, the European Parliament offers further examples of how ASR is implemented and used in institutional settings.

A first example concerns ASR used for internal documentation under tight deadlines. According to the Annual Activity Report of the European Parliament's Directorate-General for Translation, the tool Verba is used in language units to transcribe oral speeches during plenary sessions in all 24 EU languages. In 2023, it replaced the previous tool CRE because of IT constraints, and the report states that end-to-end tests produced positive results (European Parliament, Directorate-General for Translation 2024: 34).

However, although this suggests that the tool worked well enough for operational use, the report does not provide detailed evaluation data such as accuracy measures, typical error types, correction effort or feedback from the internal staff who use the tool in daily production. Verba therefore appears in public documentation primarily as a functioning internal service, but one whose practical use remains only partially visible from outside.

A second example from the European Parliament concerns ASR used to support accessibility for the public. In the same report, DG TRAD notes that the Speech-to-Text Unit completed an innovation partnership procedure to develop and acquire a platform for parliamentary events in all 24 EU languages that provides both real-time speech-to-text transcription and real-time translation. The final stage, completed in November 2023, included the addition of language models and improvements to both the ASR and machine-translation components for all 24 EU languages. The report further states that the tool is intended to support access for deaf and hard-of-hearing audiences and presents pilot findings for 2023: 45 hours of live streaming, 1,031 unique users and a user satisfaction score of 3.78/5. In addition to this score, the report notes users' overall satisfaction with the innovation provided, the ease of access to the tool, and the linguistic output of both the live transcription and the live translation, while also identifying user experience as an area for further improvement. No further direct user comments, no breakdown by language or event type, and no more detailed evidence on accuracy, error patterns or correction needs are reported (European Parliament, Directorate-General for Translation 2024: 34).

Although public institutional reporting remains limited, a public visual representation of the European Parliament's internally developed Speech-to-Text platform is available in slides from Valter Mavrič's keynote presentation at Translating and the Computer 44. The slides show both

a single-language view, in which the user can read either the original transcript or the machine translation, and a double-language view, in which the original and the machine translation are displayed side by side. In this respect, the public visual material suggests that the platform was designed as a user-facing accessibility service intended to support direct reading and multilingual access during parliamentary events (Mavrič 2022: 35, Figure 2).

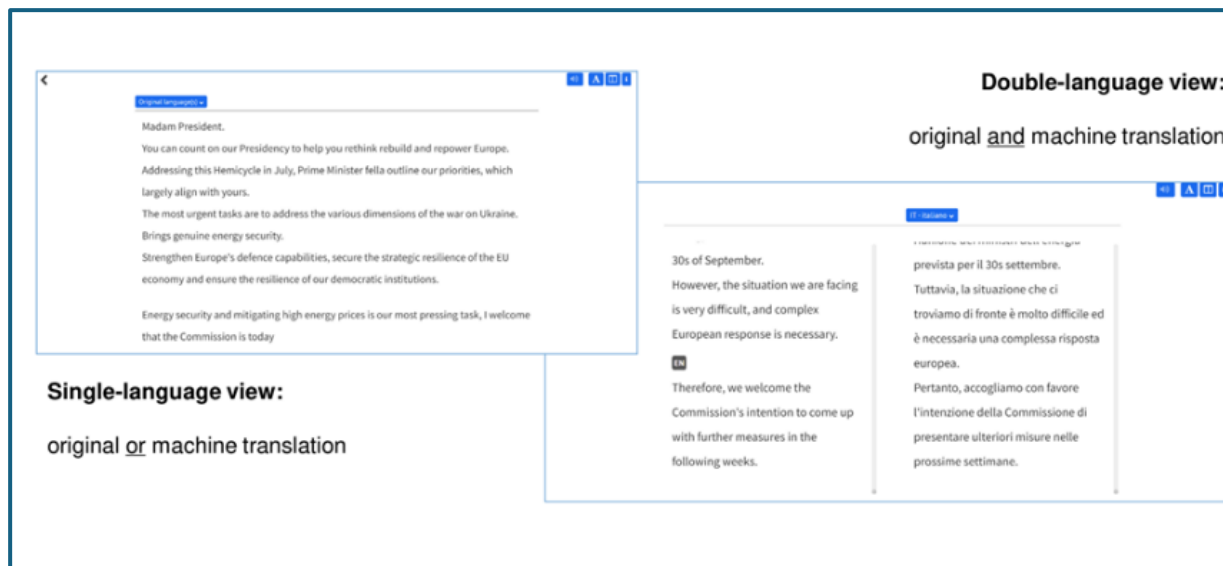


Figure 2: Interface examples of the EP's Speech-to-Text platform (Mavrič 2022: 35)

The same public presentation also shows the platform's disclaimer screen. The disclaimer states that the real-time transcription and machine translation are automatically generated, do not constitute an authentic record of the event and may be affected by the quality of remote connections. It further notes that the results are not to be considered an official transcript and that accuracy cannot be guaranteed (Mavrič 2022: 36, Figure 3).

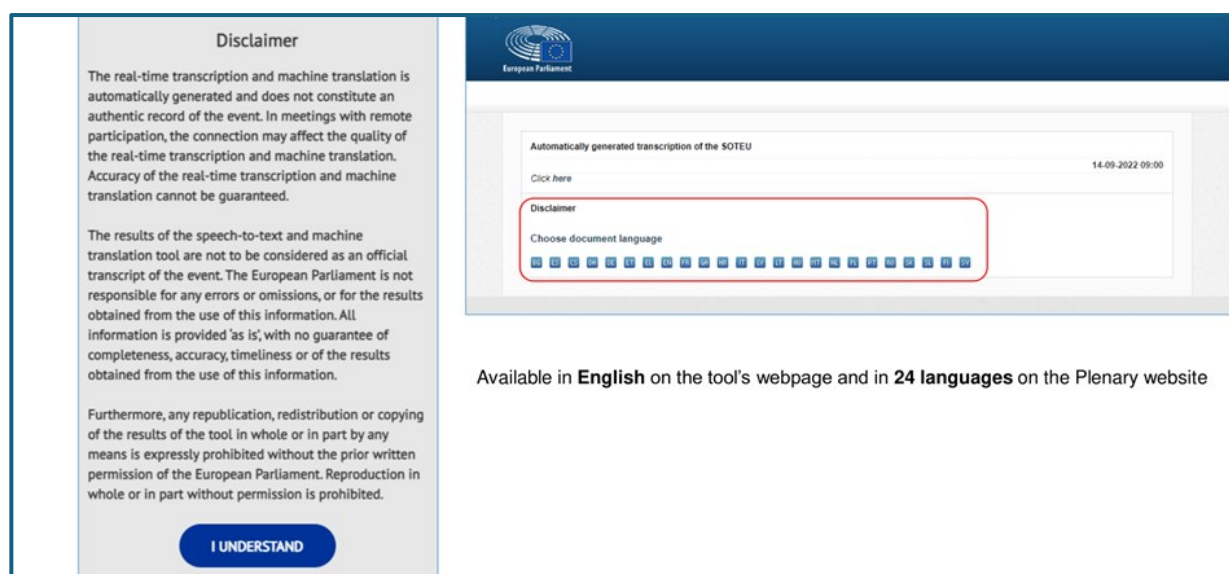


Figure 3: Disclaimer preface and language-selection screen of the EP's Speech-to-Text platform (Mavrič 2022: 36)

These important details show that the platform itself explicitly communicates its limits to users. In other words, the output is framed in advance as support for accessibility and live comprehension, not as an authoritative institutional record, in contrast to the Verba tool, mentioned above. No comparable public screenshot, interface walkthrough or other verifiable visual representation of Verba could be identified online at the time of writing. This contrast is itself further meaningful. Whereas the live Speech-to-Text platform appears, at least to some extent, as a public-facing service with a visible interface and an explicit disclaimer for users, Verba remains an internal tool used in a controlled transcription workflow. Its public invisibility is consistent with its institutional role: unlike the live platform, Verba is not intended for broad public access, but for internal use in the preparation of transcriptions that remain subject to human revision before publication.

From this perspective, the live Speech-to-Text platform can also be seen as a service whose usability depends strongly on event type and communicative setting. A satisfaction score of 3.78/5 suggests that the tool can be usable enough for real-time access and for supporting accessibility, but it also indicates room for improvement, especially if the same system is expected to perform equally well across very different communicative settings. In high-stakes contexts such as plenary speeches, where the content becomes part of public parliamentary communication and the institutional record, tolerance for errors is lower. By contrast, in more informal events (the European Parliament Youth Event or similar), users may prioritise general comprehensibility and accessibility over high quality. This distinction is relevant for the present thesis because it shows that institutional ASR must be examined in relation to use case, audience and the consequences of errors, including whether the Romanian output is perceived as sufficiently reliable for the most public and time-critical events.

Overall, the institutional documentation discussed in this subchapter shows that, across European institutions, ASR is increasingly treated not as an optional translation-workstation feature, but as multilingual infrastructure for transcription, accessibility and live institutional communication. At the same time, this institutional perspective makes several defining constraints visible.

First, multilingual ASR performance remains uneven across EU languages because it depends on the availability of in-domain speech with aligned transcripts, and this distribution is not the same for all languages. Second, institutional ASR must function within strict governance frameworks, since confidentiality obligations, data-protection rules and IT access controls shape which solutions can be deployed, how speech data can be processed and whether cloud-based services are acceptable in practice. Third, the usability and acceptability of institutional ASR also depend on the specific use of this technology. The live Speech-to-Text platform is publicly framed, through its own disclaimer, as an automatically generated aid for accessibility and live understanding, not as a guaranteed or authoritative record. Verba, by contrast, is an internal tool used in a controlled workflow where human revision remains possible before publication. This means that the threshold of acceptance differs according to function: live ASR may be accepted despite imperfections because speed and access are prioritised, while internal transcription is accepted only insofar as the output remains manageable, revisable and reliable enough for institutional use.



In this institutional landscape, where speech technologies are increasingly present, the central question is therefore not merely how ASR has advanced technically, but how it works for a specific language in a specific institutional setting and how its different use forms are experienced by professional users.

## 2.4 ASR and the Romanian language community

Building on the shift from “is ASR used in institutions?” to “how does ASR perform for a specific language?”, the aim of this subchapter is to situate Romanian first within the general ASR context and then within the broader European institutional ASR landscape.

At least until 2020, Romanian ASR was constrained by the limited availability of large, openly accessible speech corpora with reliable transcriptions. Georgescu *et al.* make this explicit: despite multiple efforts to improve Romanian speech and language resources, Romanian was still considered under-resourced, and at the time of their corpus release the largest publicly available speech corpus contained fewer than 50 hours. In their overview, they also show that Romanian speech resources tended to be measured in “hours or tens of hours”, with only a few cases reaching several hundred hours, and that not all resources were public (Georgescu *et al.* 2020, Table 2).

| Name                                | Speech Type | Domain                       | Size [hours] | Availability |
|-------------------------------------|-------------|------------------------------|--------------|--------------|
| RASC [Dumitrescu et al.2014]        | Read        | Wikipedia Articles           | 4.8          | Public       |
| RoDigits [Georgescu et al.2018]     | Read        | Spoken Digits                | 38.0         | Public       |
| RO-GRID [Kabir and Giurgiu2011]     | Read        | General                      | 6.6          | Public       |
| IIT [Bibiri et al.2013]             | Read        | Literature                   | 0.8          | Non-Public   |
| N/A [Boldea et al.1998]             | Read        | Eurom-1 Adapted Translations | 10.0         | Non-Public   |
| N/A [Popescu et al.]                | Spont.      | Internet, TV                 | 4.0          | Non-Public   |
| RSS [Stan et al.2011]               | Spont.      | Internet, TV                 | 4.0          | Public       |
| SWARA [Stan et al.2017]             | Read        | Newspapers                   | 21.0         | Public       |
| MaSS [Boito et al.2019]             | Read        | Bible                        | 23.1         | Public       |
| N/A [Tarján et al.2012]             | Spont.      | Broadcast News               | 31.0         | Non-Public   |
| N/A [Suciu et al.2017]              | Spont.      | Banking                      | 40.0         | Non-Public   |
| SSC-train1 [Cucu et al.2014]        | Spont.      | Radio and TV                 | 27.5         | Non-Public   |
| SSC-train2 [Georgescu et al.2017]   | Spont.      | Radio and TV                 | 103.0        | Non-Public   |
| SSC-train3 [Georgescu and Cucu2018] | Spont.      | Radio and TV                 | 49.5         | Non-Public   |
| SSC-train4 [Georgescu et al.2019]   | Spont.      | Radio and TV                 | 280.0        | Non-Public   |
| CoRoLa [Mititelu et al.2014]        | N/A         | Various Sources              | 152.0        | Cvazi-Public |

Table 2: Overview of the availability of Romanian speech corpora (Georgescu et al. 2020: 6607)

ASR performance for a specific language depends on resource conditions: how much transcribed speech exists, how diverse the speakers are, how representative the recording conditions are and how well the data match the target domain. A central line of Romanian ASR research has therefore focused on building speech corpora. As a response to this lack, Georgescu *et al.* contribute a corpus of 100 hours of read speech from 164 speakers (Georgescu *et al.* 2020: 6606–6607).

This picture is further updated by more recent open-data initiatives. In particular, the Romanian release of Mozilla Common Voice Scripted Speech v25.0, published in March 2026, reports 44,420 clips, representing 49.92 hours of recorded speech, including 22.3 validated hours, from 461 speakers (Mozilla Common Voice 2026)<sup>6</sup>. This does not remove the broader resource gap for Romanian ASR in domain-specific or institutional settings, but it does show that openly available Romanian speech data have continued to expand since the overview presented by Georgescu *et al.* (2020). At the same time, the comparison suggests that recent progress lies not only in the number of hours, but also in broader community participation and speaker diversity.

The importance of such corpora lies in creating reproducible benchmarks, enabling comparability across systems and lowering the entry barrier for research. The question that remains, however, is to what extent Romanian ASR has moved from research corpora into widely available tools. At the level of general operating-system tools, Microsoft Windows voice typing supports Romanian, as mentioned among supported languages on its website (Microsoft Windows n.d.)<sup>7</sup>. In translation-tool environments, the dictation feature in Trados Online Editor also supports Romanian (RWS n.d.)<sup>8</sup>. However, one of the best-known desktop dictation products, Dragon Professional, formerly known as Dragon NaturallySpeaking, is officially available only in English, German, French, Spanish, Dutch and Italian. Romanian is not included (Nuance 2023)<sup>9</sup>.

Another recurring finding in Romanian ASR research is that recognition errors are not caused only by acoustic issues, but also by the quality of the written Romanian used to train the language model (orthographic issues). Avram *et al.* illustrate this in their end-to-end Romanian ASR study based on DeepSpeech 2, an end-to-end neural approach in which a model learns to map audio directly to text, reducing reliance on a traditional ASR pipeline composed of separate hand-built components (pronunciation dictionaries and manually designed acoustic models). They train an end-to-end ASR system and compile 230 hours of aligned Romanian speech for training. In addition, they train a separate language model intended to reduce recognition errors. For this language model, they use two large text sources: CoRoLa (reported as 941,204,169 tokens) and the Romanian portion of OSCAR (reported as circa 11 GB). However, they emphasise that the language model is highly sensitive to text noise and therefore, they apply extensive cleaning, including removing lines without diacritics or with incorrect diacritic encoding and applying Romanian-specific character normalisation – replacing “ı” with “â” in certain contexts (Avram *et al.* 2020: 397–398).

A second Romanian-specific writing challenge concerns hyphenation and clitic constructions, where what is spoken and what is expected in standard written Romanian can diverge in ways that increase correction effort. Avram *et al.* note that raw ASR output does not insert hyphens, so forms such as *s-au* may appear as *sau*. Both are valid forms, so choosing the correct variant requires the surrounding context. This is precisely the type of ambiguity where language modelling can help, because it uses context to favour more plausible word sequences. In their system, Avram, Păiș and Tufiș describe the language model as a component that

<sup>6</sup> <https://mozilladatacollective.com/datasets/cmn2e8rmi0116mm07vxurptse>

<sup>7</sup> <https://support.microsoft.com/en-us/windows/use-voice-typing-to-talk-instead-of-type-on-your-pc>

<sup>8</sup> <https://docs.rws.com/trados-enterprise-%26-accelerate-791611/dictation-in-trados-online-editor-1123284>

<sup>9</sup> [https://nuance.custhelp.com/app/answers/detail/a\\_id/6280/~/what-languages-are-available-for-dragon](https://nuance.custhelp.com/app/answers/detail/a_id/6280/~/what-languages-are-available-for-dragon)



supports decoding by assigning higher probabilities to more likely words given the preceding words, thus using contextual information (Avram *et al.* 2020: 399).

Overall, Romanian ASR research is moving toward larger and more diverse training setups and increasingly relies on language modelling. However, the key open issue for professional adoption is not only whether such errors can be reduced technically, but also how much they still matter in practice: how frequently they occur in domain-specific tasks, how much correction effort they add during revision and whether repeated small errors, such as problems with diacritics, hyphenation and clitic forms, reduce users' trust in ASR output and therefore its acceptance in professional Romanian-language workflows.

In addition, Romanian ASR has gained practical relevance through European institutional deployments. As discussed in subchapter 2.3., the European Parliament reports the development and acquisition of a tool for real-time speech-to-text (and machine translation) of parliamentary debates in all 24 official EU languages (European Parliament, Directorate-General for Translation 2024: 34). The European Commission likewise presents its Speech-to-Text service as available in all EU official languages, therefore including Romanian among the supported languages (European Commission n.d.a). Romanian coverage is also documented more directly in published descriptions of the European Parliament's Innovation Partnership prototype: by November 2022, Romanian is explicitly listed among the 19 language models available in the demo interface, and it is also one of the six languages selected for more detailed human evaluation (EN, FR, ES, IT, RO, DE) (Di Nuovo 2023: 185–195). Finally, European Parliament debates have also been used to build publicly released speech–text resources that include Romanian: Europarl-ST, “a Multilingual Speech Translation Corpus, that contains paired audio-text samples for Speech Translation, constructed using the debates carried out in the European Parliament in the period between 2008 and 2012” (Europarl-ST n.d.). These developments are important bridges for Romanian to institutional ASR contexts.

The literature in this subchapter has shown that Romanian ASR has previously been shaped by resource constraints, with progress driven by corpus building and by efforts to improve robustness and written-output quality (especially diacritics consistency, hyphenation and other orthography-dependent forms). At the same time, Romanian has become practically relevant beyond research, because it is now included in European institutional ASR deployments and in institutional resources derived from European Parliament data. However, the literature still leaves an important gap: most Romanian ASR studies assess progress mainly in technical terms or in general professional contexts, while institutional documentation tends to confirm language coverage without providing Romanian-specific evidence on usability in institutional work.

This thesis aims to contribute to the current state of research on Romanian ASR by examining its use in a concrete European institutional setting, the European Parliament, and by focusing on user-centred criteria: relevance in relation to workflow fit, usability in relation to correction effort and acceptance under institutional constraints. Its main purpose is to complement existing Romanian ASR research and to provide evidence that can be directly useful for improving the deployment and integration of Romanian ASR in European institutional environments.

## 2.5 Future directions in ASR: rationale and research gaps

Taken together, the previous subchapters show that automatic speech recognition can no longer be understood as an emerging technology or as a purely technical field of research. As discussed in subchapter 2.1, ASR has developed from early, tightly controlled recognition tasks into largely speaker-independent, large-vocabulary systems capable of supporting both speech-to-text functions, such as dictation and transcription, and speech-enabled interaction, such as voice commands (Juang & Rabiner 2004: 3–18; Furui 2005; Yu & Deng 2015: 1–3).

Recent progress has been driven not only by increased computing power and larger datasets, but also by data-driven and AI-based approaches designed to improve performance under more realistic conditions, including spontaneous speech, disfluencies, background noise and pronunciation variability (Kheddar *et al.* 2024). In this sense, ASR has become an everyday technology and, at the same time, a potentially relevant tool for professional language work.

At the same time, the literature reviewed in subchapter 2.2 makes clear that technical capability does not automatically translate into professional uptake, especially in translation workflows. In professional translation environments, and particularly in institutional ones, the decisive question is not simply whether ASR technically works, but whether it fits the workflow in which it is meant to be used. Such fit depends on whether ASR is experienced as relevant to the actual task, for example for drafting or revising in a CAT tool, usable with a manageable correction effort, and acceptable under organisational constraints such as confidentiality, data protection and IT governance. In this sense, ASR in translation should be assessed not in abstract technical terms alone, but in relation to its actual workflow fit: whether it supports concrete translation and revision tasks efficiently, reduces rather than increases correction effort and remains compatible with the institutional conditions under which professional translators work (Ciobanu & Secară 2019: 96–101; Zapata & Søeborg Kirkedal 2015: 207–208; Zapata *et al.* 2023: 16–17).

Subchapter 2.3 further shows that, although translators may adopt ASR individually as a productivity or ergonomics strategy, European institutions increasingly use ASR at the level of infrastructure, particularly for multilingual communication, accessibility and time-critical language mediation. This means that institutional value can exist even where workstation-level adoption remains uneven.

At EU level, the European Commission presents ASR as part of Europe’s critical digital capacities and links it to capacity-building measures such as mapping existing solutions, developing reusable prototypes and collecting language data. Two points are especially important in this context. First, progress depends on the availability of in-domain speech with aligned transcripts, and this availability remains uneven across languages. Second, institutional ASR services must comply with governance requirements, especially regarding where speech data are processed, how they are stored and who can access them. For this reason, institutions tend to favour controlled environments and internal services over unregulated public tools. The same institutional logic is visible in the European Language Data Space, which promotes the secure sharing and reuse of language data and related resources under EU governance in order to strengthen multilingual AI tools, including ASR (European Commission 2024; LDS n.d.; LDS 2025; Wang *et al.* 2021: 994–995).

At the level of the European Parliament, institutional ASR is reflected in tools such as Verba, used for the transcription of plenary speeches, and the Speech-to-Text Unit's live ASR platform, used to support accessibility during parliamentary events. However, the available institutional reporting often confirms only that such tools exist, without offering sufficient detail on how they are developed, adapted and evaluated in practice. For example, public information on the Parliament's live Speech-to-Text platform does not explain in concrete terms how it performs in different settings, how it has been developed operationally, or in which situations live ASR is preferred over other forms of linguistic mediation. Likewise, the available information on Verba does not include feedback from its users and does not provide language-specific evidence, including for Romanian (European Parliament, Directorate-General for Translation 2024: 34; Mavrič 2022: 35–36).

From a workflow perspective, European institutions also provide little public information on active ASR use in everyday professional tasks such as drafting emails, taking notes or transcribing online meetings on institutional devices, nor on how ASR may be integrated into the CAT-tool environment currently used in the European Parliament for translation and revision.

Subchapter 2.4 situates Romanian within these broader developments and shows that Romanian ASR has been shaped both by resource limitations and by language-specific output-quality challenges. Romanian has historically been described as under-resourced in terms of large, openly accessible speech corpora with reliable transcriptions, which explains why corpus building and benchmark creation have become central directions in Romanian ASR research (Georgescu et al. 2020: 6606–6607).

At the same time, an important question remains: to what extent has Romanian ASR moved from research corpora into widely available tools? Romanian is supported in some general and translation-related environments, such as Windows voice typing and Trados Online Editor, but it is still absent from one of the best-known desktop dictation products, Dragon Professional (Microsoft Windows n.d.; RWS n.d.; Nuance 2023).

In addition, Romanian ASR studies show that recognition quality is influenced not only by acoustic conditions, but also by orthographic and language-modelling issues, especially in relation to diacritics, hyphenation and clitic constructions. Such recurring issues may increase correction effort and affect perceived usability (Avram et al. 2020: 397–399). Romanian is also increasingly present in institutional ASR contexts and resources derived from European Parliament data, yet an important gap remains: institutional documentation tends to confirm language coverage without providing Romanian-specific evidence on usability and acceptance in concrete professional settings.

Overall, ASR emerges as a technology whose value depends not only on technical performance, but also on how it is shaped by concrete linguistic conditions, the demands of professional workflows and the institutional environments in which it is deployed. ASR cannot be evaluated through a single measure of success. In live institutional settings, speed, continuity and general comprehensibility may be prioritised; in internal transcription workflows, revisability and manageable post-correction effort become more important; and in office or CAT-tool environments, the central issue is whether speech input actually improves workflow

efficiency without creating additional effort. A system may therefore be acceptable in one institutional function while remaining marginal in another.

Against this background, several research gaps become particularly visible. First, there is still limited transparency regarding how current institutional ASR solutions in European institutions are developed, adapted and evaluated in practice. Second, existing research has not sufficiently examined ASR as a workflow-dependent technology whose value may vary across different forms of use, including live transcription, asynchronous transcription and active dictation or voice-enabled interaction in office and CAT-tool environments. Third, there is a specific lack of Romanian-focused evidence at the intersection of ASR, institutional use and professional user experience. More precisely, the literature does not yet provide a sufficiently detailed account of how Romanian-language professionals in European institutions perceive the relevance, usability and acceptance of currently available ASR technologies under actual organisational conditions. These observations underline the need to examine ASR more closely in relation to concrete institutional workflows, language-specific conditions and professional user experience.

### **3. Research questions and methodology**

Across the reviewed literature and institutional documentation, several gaps emerge which this thesis aims to address through the following research questions:

#### ***Limited transparency on how institutional ASR is developed and perceived in practice***

RQ1: How are current institutional ASR solutions in the European Parliament described and evaluated by experienced professionals?

#### ***From technical capability to institutional workflow suitability in translation***

RQ2: How can ASR be integrated into workflows in the European Parliament, and more specifically into the CAT-tool workflow, under organisational constraints?

#### ***Romanian-specific ASR insights***

RQ3: How do Romanian-language professionals in the European Parliament perceive and experience ASR technologies?

To answer these questions, the thesis adopts a mixed-methods design based on empirical data collected within the European Parliament, with a particular focus on Romanian-language workflows. In line with the analytical framework developed in chapter 2, the study distinguishes between three forms of ASR use, because they serve different functions and involve different expectations in terms of relevance, usability and acceptance:

- **Live / synchronous ASR** for real-time transcription, where speech is transcribed in real time to support accessibility and time-critical communication during events (public use of ASR);
- **Asynchronous ASR** for non-live transcription, where speech is converted into text primarily for documentation and internal processing and users work mainly with the generated output (transcription workflow using Verba);
- **Active ASR use** for dictation and voice-enabled interaction, where users speak to produce text and may use voice-enabled commands to support drafting, translating or revising in office applications and in a CAT-tool environment.

These three forms of ASR use structure the study both analytically and methodologically. The empirical part is organised in the same order, moving from live / synchronous ASR to asynchronous ASR and finally to active ASR use. At the methodological level, the thesis combines interviews, a presentation, a follow-up try-out and a questionnaire. The first two empirical steps are based on the same method, namely interviews, but they address different forms of ASR use and different perspectives. The subsequent steps differ methodologically, since they are designed not only to demonstrate active ASR use in practice, but also to prepare participants for the questionnaire that follows. The sequence of methods also serves a triangulating purpose. The two interviews anchor the study in already existing institutional ASR practices, whereas the presentation, try-out and questionnaire extend the analysis to possible future workstation-level uses. This progression from deployed solutions to potentially adoptable ones makes it possible to compare not only what is operationally available in the European Parliament, but also what appears professionally desirable yet institutionally unavailable.

First, a semi-structured interview was conducted with Dragoş-Lucian Ciobanu, Head of the European Parliament's Speech-to-Text Unit, in order to gather detailed information on the development, use and evaluation of the Parliament's live ASR platform, with particular attention to Romanian. This interview addresses live / synchronous ASR for real-time transcription, since it focuses on a service used in public and time-critical settings to support accessibility. It contributes primarily to RQ1 by providing an expert perspective on how current institutional ASR solutions in the European Parliament are described and evaluated in terms of relevance, usability and acceptance.

Second, a short interview was conducted with an anonymous experienced Romanian proofreader from the European Parliament Romanian Unit, who has extensive experience using the internally developed Verba tool in daily work. Although this empirical step is methodologically the same as the first one, namely an interview, it addresses a different form of ASR use. More specifically, it examines asynchronous ASR for non-live transcription, since it focuses on a tool used for the documentation and internal processing of parliamentary speeches. It contributes to RQ1 by complementing the expert interview with an experienced user perspective on another current institutional ASR solution in the European Parliament, namely Verba. At the same time, it also contributes to RQ3 by providing the perspective of a

Romanian-language professional working directly with Romanian ASR output in an institutional setting.

The third form of ASR use, active ASR for dictation and voice-enabled interaction, is approached through a different empirical sequence. As a first step, ASR solutions available on European Parliament laptops were identified and tested in order to prepare a structured presentation for the European Parliament Romanian Unit. The presentation introduced available ASR options for everyday tasks, such as dictating in Word or drafting emails, and also discussed possible uses of ASR in the translation process, including in Trados Studio. In this way, it contributes primarily to RQ2 by exploring how ASR may be integrated into workflows in the European Parliament, and more specifically into the CAT-tool workflow, under organisational constraints. In a further step, the presentation was followed by an optional try-out session. This try-out had a preparatory function within the research design, since it allowed participants to engage more directly with the demonstrated tools and to relate them to their own professional practice before responding to the questionnaire. It therefore formed an intermediate step between the presentation itself and the collection of participant feedback. Finally, a questionnaire was used to gather participants' impressions of the demonstrated ASR tools and of their possible relevance for Romanian-language work in the European Parliament. Building on the presentation and the follow-up try-out, the questionnaire contributes primarily to RQ3 by documenting how Romanian-language professionals perceive these technologies in relation to their own work practices, while also supporting RQ2 insofar as it provides evidence of their perceived suitability for workflow integration.

Taken together, this design makes it possible to examine the three forms of ASR use through complementary empirical steps. The interviews provide two perspectives on existing institutional ASR solutions, namely an expert perspective on live ASR and a user perspective on asynchronous ASR. The presentation, follow-up try-out and questionnaire, in turn, address active ASR use in a more practice-oriented way, moving from demonstration to direct engagement and then to structured feedback. In this way, the thesis analyses not only what ASR solutions are currently available in the European Parliament, but also how they are experienced and assessed under real institutional conditions.

### **3.1 Interview with the Head of the European Parliament Speech-to-Text Unit**

To examine live ASR for real-time transcription in the European Parliament and to complement the limited information available in institutional reporting, this thesis includes a semi-structured interview with Dragoş-Lucian Ciobanu, Head of the Speech-to-Text Unit. The purpose of this interview was to document how live ASR is developed, used and evaluated in practice, with particular attention to Romanian and to the institutional conditions that shape relevance, usability and acceptance.

Semi-structured interviewing was chosen because it combines a clear thematic structure with sufficient flexibility for detailed explanation. Following Saldanha and O'Brien (2013), the interview was guided by a predefined set of questions while still allowing the interviewee to elaborate on experiences, clarify institutional procedures and introduce relevant aspects not fully captured by predetermined questions. This format was particularly appropriate for the

present study because institutional ASR solutions are often described only at service level in public documents, without providing insight into operational decision-making, evaluation practices or user-feedback mechanisms.

The interview was conducted in Romanian via Microsoft Teams and recorded with the participant's informed consent. The recordings were then transcribed for analysis and translated into English for inclusion in the thesis. In line with Mayring (2022), the transcription focused on content relevant to the research questions, while material not directly relevant to the analysis was omitted. Following Dresing and Pehl (2015), a simplified transcription method was used, since the analytical focus lies on content rather than on fine-grained linguistic or interactional detail. For this reason, non-verbal elements and hesitations were not represented systematically. The full documentation of the interview is provided in Appendix A, which includes the bilingual interview questions (English and Romanian), the revised transcripts (English and Romanian) and the corresponding recordings (Romanian).

The interview guide was organised into four thematic blocks:

#### Block 1: General questions

This first block establishes the general institutional context by clarifying the interviewee's role and the responsibilities of the Speech-to-Text Unit, so that the later answers can be interpreted with a clear understanding of what the unit is responsible for and what falls within its scope. It also clarifies how the Speech-to-Text Unit positions ASR within the Parliament, while questions on advantages and challenges are intended to capture both the perceived value of live ASR and the main factors that shape its use in the Parliament, such as quality requirements, operational constraints and governance-related issues.

#### Block 2: Questions on the internally developed Speech-to-Text platform

This second block goes beyond the general descriptions found in institutional documents by asking for practical details on how the platform was developed and adapted for parliamentary needs. A terminology-focused question targets a key issue in institutional language work, since parliamentary discourse includes specialised terminology, acronyms and proper names and it also helps assess whether domain adaptation supports quality in practice. Further, a question regarding benchmarking focuses on how the platform is evaluated and how automatic live transcription compares to a respeaking workflow, where a human mediator can reduce errors.

#### Block 3: Romanian-focused questions

This third block addresses the Romanian-specific gap discussed in the theoretical foundation, since institutional documentation often confirms language coverage without providing Romanian-specific information on performance or user experience. The questions aim to clarify the status of Romanian ASR support and whether Romanian is treated as fully operational or still improving. Another second question invites discussion of Romanian-specific factors that may influence recognition quality and correction effort, linked directly to usability and to users' trust in the ASR output.

#### Block 4: Forward-looking question

This final block is intended to capture a forward-looking institutional perspective. It allows the interviewee to highlight which limitations matter most at present and which improvements would be most useful for the further integration of ASR within the European Parliament. It also supports the future directions discussion by identifying priorities that could increase relevance, usability and acceptance, such as better data, improved evaluation methods, stronger domain adaptation, clearer integration into workflows or more language-specific development.

Overall, this interview guide moves from the general institutional context to platform-specific implementation details and finally to Romanian-specific details and future priorities regarding ASR. The interview findings are reported in subchapter 4.1 and are interpreted in relation to the three analytical dimensions of this thesis, namely relevance, usability and acceptance.

### **3.2 Interview on Verba with an experienced user from the European Parliament Romanian Unit (DG TRAD)**

Institutional ASR in the European Parliament is not limited to live transcription for accessibility but also includes an internal tool that supports documentation workflows. To capture a user perspective on asynchronous ASR, this thesis includes a short interview with an anonymous experienced Romanian proofreader from the European Parliament Romanian Unit, who uses Verba regularly, an internally developed tool deployed within the Directorate-General for Translation to transcribe speeches during plenary sessions.

The purpose of this interview was to collect practice-based information on how Verba is used in daily work and how it is perceived by an experienced Romanian professional user in terms of relevance, usability and acceptance. In contrast to the expert interview with the Head of the Speech-to-Text Unit, which addresses development and institutional evaluation, this interview focuses on the user perspective. More specifically, it examines the experience of working with automatically generated Romanian transcriptions, the typical correction effort involved, and whether Verba is perceived as supporting efficiency under time pressure. The interview also makes it possible to identify recurring Romanian-specific issues that may affect usability, such as the handling of terminology, acronyms and proper names in parliamentary discourse.

The interview was conducted in person in Romanian and recorded with the participant's informed consent. Because Verba is intended for internal European Parliament use, the interview was deliberately kept brief and limited to general aspects of relevance, usability and acceptance, without discussing sensitive internal details or providing a technical evaluation of the tool. The participant also requested anonymity, and the interview is therefore presented in anonymised form. The recording was transcribed and translated into English for analysis. As in the first interview, transcription was content-focused and selective, following Mayring (2022), and a simplified transcription approach was applied in line with Dresing and Pehl (2015). The full documentation of this interview is provided in Appendix B, which contains the interview



questions, the revised Romanian and English transcripts, and the corresponding Romanian recording.

The interview questions were designed to connect everyday practice with user feedback on an asynchronous transcription workflow. The first question established the practical context by asking what Verba is and how it is used in the interviewee's daily work as a Romanian proofreader at the European Parliament. This made it possible to situate the tool within the broader workflow before moving to evaluation. The second question focused on perceived efficiency and aimed to clarify whether Verba is experienced as making transcription work faster and more manageable under real production conditions. The third question addressed perceived accuracy in the specific setting of plenary sessions, where speech rate, speaker variability and dense institutional content may affect output quality. The fourth question turned to post-correction effort and typical error patterns, thereby providing a more concrete indicator of usability and helping to identify which kinds of mistakes remain relevant in Romanian.

Taken together, these questions make it possible to approach Verba not as an abstract ASR solution, but as a tool embedded in the daily practice of Romanian-language transcription at the European Parliament. In this way, the interview contributes a practice-based and user-centred perspective on asynchronous ASR under concrete institutional conditions.

### **3.3 Presentation for the European Parliament Romanian Unit (DG TRAD)**

In order to examine active ASR use for dictation and voice-enabled interaction in an institutional context, a structured presentation was delivered in Romanian during a meeting of the European Parliament Romanian Unit. The presentation was given by the author in her capacity as a proofreader-language editor in the unit and constituted the first step in the empirical sequence used to examine active ASR use. Instead of using a conventional PowerPoint presentation, a pre-recorded video was prepared so that the demonstrated functions could be shown step by step in their actual digital environments.

The presentation was based on two types of material. First, it included functions that were demonstrated directly in the video. Second, it drew on information that was publicly available at the time of the research through official institutional pages, software documentation and support pages. Accordingly, the methodology section reports only the demonstrated content and the publicly documented features on which the presentation was based. By contrast, institution-specific issues that emerged only during the preparation and testing of the presentation, such as internal technical restrictions or communication with support IT services, are discussed separately in subchapter 4.3 as findings.

At the beginning of the presentation, the topic of the thesis was briefly introduced and the research aim was explained, namely to examine the relevance, usability and acceptance of ASR in European institutional contexts, with particular attention to Romanian. In order to make the topic more accessible, the presentation opened with a familiar everyday example: the use of voice assistants and speech input while driving. A short video recorded in a car was used to illustrate how speech can serve as an input method in a hands-free situation, for example by dictating a destination to Siri. This introductory example positioned ASR first as an everyday

technology and only afterwards as a professional and institutional tool, thereby making the later transition to office applications and CAT tools easier to follow. At the same time, it reflected the broader distinction developed in the theoretical foundation between ASR as speech-to-text and ASR as human–machine interaction.

The next part of the presentation introduced ASR more explicitly at a conceptual level. The explanation briefly outlined how speech recognition works in general terms: speech is captured as an audio signal, acoustic patterns are processed, the input is interpreted through linguistic modelling, and the result is then decoded either into written text or into a voice-triggered action. The explanation was kept short and non-technical, but it was intended to establish a common minimum understanding before the demonstrations began.

After this general introduction, the presentation moved to the institutional level by referring to the European Parliament’s Speech-to-Text platform as an example of live ASR used for accessibility and multilingual communication. As discussed in subchapter 2.3, public institutional documentation states that DG TRAD’s Speech-to-Text tool has expanded its coverage to all 24 official EU languages and can automatically transcribe and translate multilingual parliamentary debates in real time (European Parliament, Directorate-General for Translation 2024: 34). In the presentation, this example was used to show that ASR in European institutions is not only a private productivity tool, but also part of a broader service infrastructure that supports accessibility and participation.

A central practical section of the presentation focused on Windows Voice Typing as a basic example of desktop dictation. As outlined in subchapter 2.3, this built-in speech-to-text function can normally be used in many desktop applications for tasks such as drafting emails or taking notes. According to Microsoft Support, it can be activated with the keyboard shortcut Windows key + H, and Romanian is included among the supported Windows 11 voice typing languages (Microsoft Windows n.d.). In the presentation, the function was demonstrated first in an email window and then in Word, in order to show how speech-to-text can be used for simple tasks such as drafting a short email and taking notes (Figure 4).

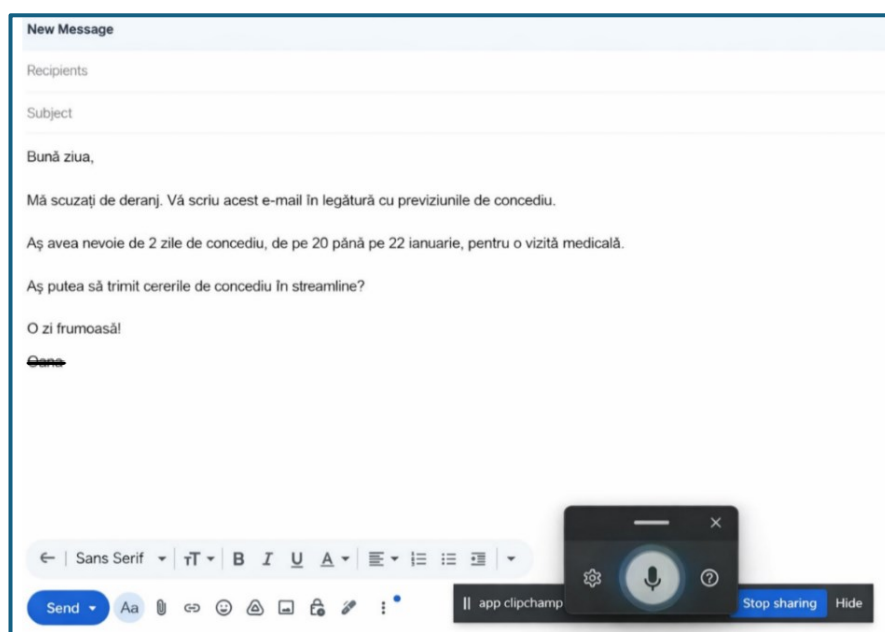


Figure 4: Screenshot from the presentation video (Appendix C) showing an email drafted in Romanian with Windows 11 voice typing

This part of the presentation served as a basic example of ASR use outside a CAT tool. At the same time, it also illustrated in practical terms the institutional limitation discussed in subchapter 2.3. On EP-managed laptops, Windows voice typing was not available. When the function was activated on such a device, a system notification appeared stating that the speech service was managed by the organisation and that the feature was not available (Figure 1). This example was included in the presentation because it showed clearly that a widely available ASR function may exist in principle but still remain unavailable in the institutional environment of the European Parliament.

The next step in the presentation introduced Dragon Professional (also known as Dragon NaturallySpeaking) as a separate desktop speech-recognition application. As shown in subchapter 2.4, Romanian is supported by some general and translation-related tools, such as Microsoft Voice Typing in Windows 11 and the dictation feature in Trados Online Editor, but it is not included among the supported languages of Dragon Professional. Against this background, Dragon was included in the presentation as an example of a well-known desktop dictation product that offers professional ASR functions but does not currently support Romanian. Public Nuance documentation describes Dragon Professional as software that supports both front-end dictation and back-end transcription from pre-existing audio files (Nuance 2023). In the presentation, Dragon was therefore shown as software that must be installed separately and that offers not only live dictation, but also the possibility of uploading and transcribing audio files. This example was included in order to show that professional ASR is not limited to built-in Windows tools and that dedicated speech-recognition software may offer additional functions, such as audio-file transcription and more advanced command use. At the same time, it helped position Romanian more clearly within the current limits of mainstream desktop dictation software.

The presentation then moved to the translation environment. For the translation-related demonstrations, the same bilingual source material: the summary of Petition No 0886/2025 on the regulation of competitions for animal traction in the European Union, reproduced in Appendix C (European Parliament 2025). This text was selected because it provided an institutional and publicly accessible example with official English and Romanian versions, clear segmentable content, and a combination of proper names, legal references, numbers and domain-specific terminology.

The presentation focused first on the desktop version of Trados Studio. The purpose of this part was to show that a general dictation tool such as Dragon Professional can also be used inside a CAT-tool interface. In the demonstration, Dragon Professional was used in Trados Studio desktop in order to illustrate how speech input could support translation work at segment level. Using the petition text mentioned above, one Romanian source segment was opened in Trados Studio desktop and the English target segment was dictated with Dragon Professional (Figure 5).

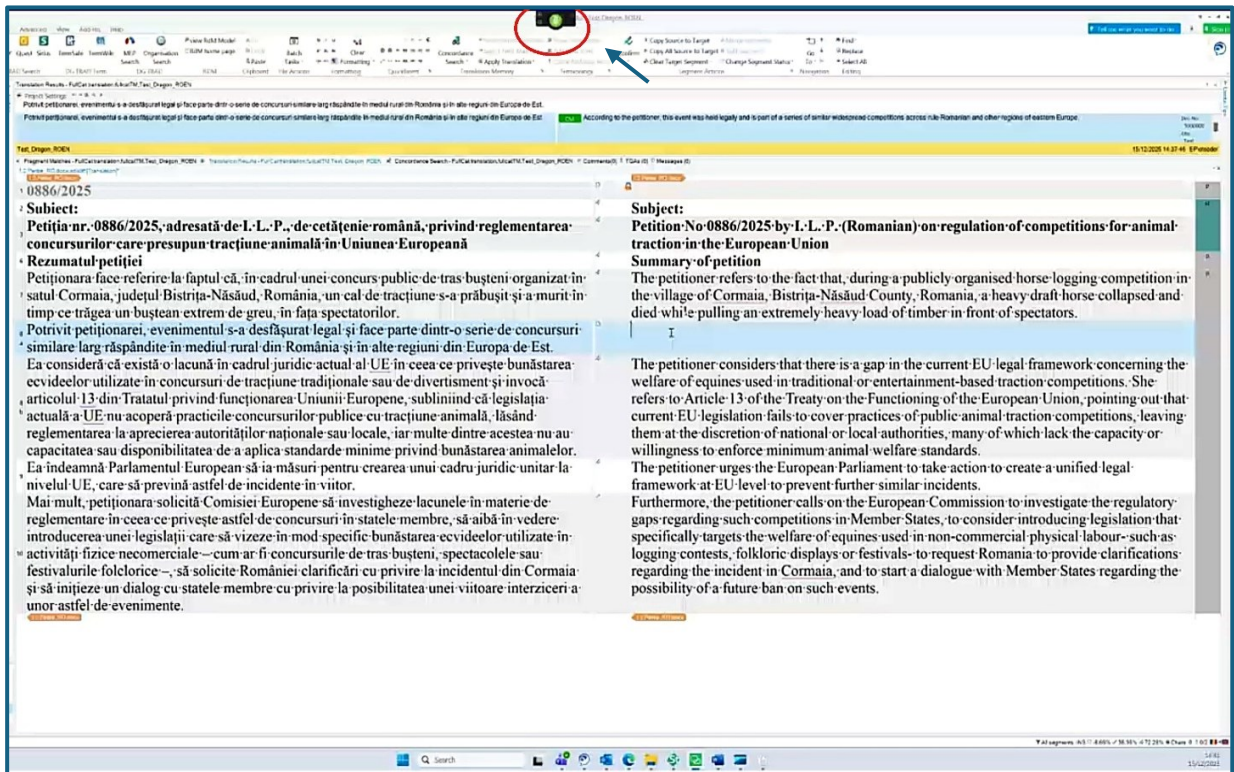


Figure 5: Screenshot from the presentation video (Appendix C) showing Dragon Professional within Trados Studio desktop

This demonstration did not aim to reproduce the full institutional workflow, especially since Romanian is not currently supported by Dragon Professional, but rather to show in functional terms that externally installed ASR software can feed spoken input directly into a CAT-tool segment. This example also reflected the discussion in subchapter 2.2, according to which ASR in translation should be examined not only as general dictation, but in relation to real workflows shaped by CAT-tool interfaces and segment-based work.

The most detailed part of the presentation focused on Trados Online Editor. According to RWS documentation, this environment includes dictation, supports Romanian as a target language, and provides documented Romanian voice commands (RWS n.d.). For this reason, it was presented as the most relevant environment for examining ASR in a CAT-tool setting and as the clearest publicly documented example of ASR integrated directly into a translation interface with Romanian-language support.

Using the same petition text as in the Dragon demonstration, the Trados Online Editor example was configured from English into Romanian. The demonstration was designed to show two different forms of ASR use within the same CAT-tool environment. First, one segment was dictated completely in Romanian from scratch, in order to illustrate full segment-level dictation. Second, other segments were left in their pretranslated form and were then revised through a combination of Romanian voice commands and manual intervention. In this way, the demonstration did not present ASR only as a drafting tool, but also as a possible aid for revision and post-editing.

During this part of the video (Figure 6), the command list was shown and several Romanian commands were illustrated, including commands for inserting punctuation, deleting text, selecting the next word, deleting the selected word, moving to the next segment, confirming a segment and moving on, and locking or blocking a segment. The video also showed that dictation was deliberately produced with short pauses and hesitations in order to approximate a more realistic translation situation in which the translator formulates the target text while thinking through the source segment. At the same time, the keyboard was used selectively whenever typing seemed more efficient than speaking.

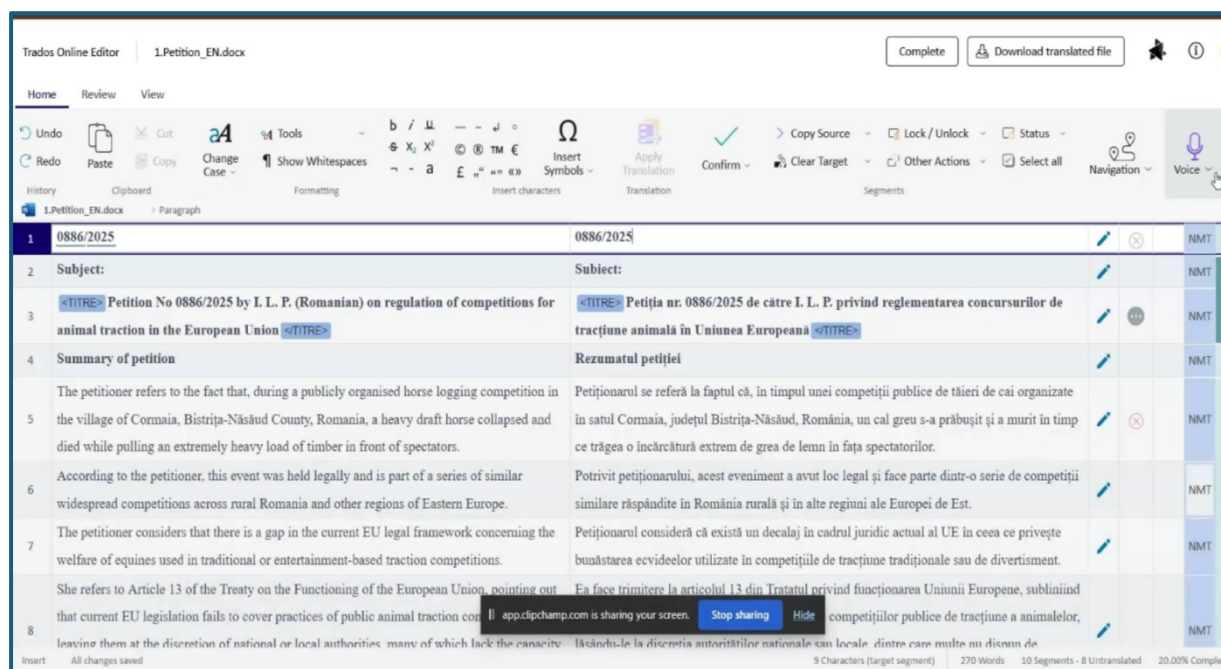


Figure 6: Screenshot from the presentation video (Appendix C) showing Trados Online Editor with Romanian dictation and voice commands

This part of the presentation reflected the discussion in subchapter 2.2 particularly clearly. The Trados Online Editor demonstration showed that ASR in translation is not necessarily a matter of fully hands-free first-draft production. Rather, it can function in a hybrid way, combining speech, keyboard input and CAT-tool commands. It also showed more concretely that the relevance of ASR in translation may lie not only in dictating an entire segment from scratch, but also in revising, correcting and navigating pretranslated material.

In this sense, Trados Online Editor was used both as a practical test case of currently available Romanian-language functionality and as a possible model for future development, with the expectation that similar features, including more specialised voice commands, could eventually become available in the institutional desktop workflow as well.

Finally, the presentation concluded with an invitation to continue the discussion in a practical way through an optional try-out session and a short questionnaire. The presentation itself did not aim to measure participants' reactions yet, but to provide a shared basis of information and a shared visual reference for the next empirical step. It introduced a sequence of increasingly specific ASR environments, from everyday mobile use to institutional live use, then to general desktop dictation and finally to CAT-tool use with Trados. In this way, the presentation created the conditions for a more informed follow-up questionnaire. The full documentation of this

empirical step is provided in Appendix C, which contains the presentation video, the text used and the explanation transcript.

### **3.4 Follow-up try-out (EP RO Unit, DG TRAD)**

Following the presentation, participants were invited to take part in an optional follow-up try-out of some of the demonstrated ASR solutions. Five respondents participated in this session and tested selected tools more practically together with the author of this thesis.

In particular, the dictation function in Trados Online Editor was used on a personal laptop within a project created specifically for the try-out session on the basis of the same bilingual petition text used in the presentation, while Dragon Professional in English was also tested on an EP laptop. The session did not focus only on general tasks such as drafting a short email (with Dragon Professional) but also included simple translation-related activities in Trados projects (within Trados Online Editor) prepared in advance from this material, including segment-level dictation, navigation, revision and the correction of pretranslated segments by voice.

Because availability was limited, not all participants were able to attend the try-out session. Moreover, those who did attend did not carry out a full translation or post-editing task with ASR in either the desktop or the online environment. Instead, the session remained exploratory and concentrated on a small number of representative actions. These included dictating an entire segment from scratch, moving between segments through voice commands and revising pretranslated segments by using commands for punctuation and deletion.

In some cases, participants also corrected text by selecting words or phrases with the mouse and then deleting or modifying them by voice rather than by keyboard. This reflects the discussion in subchapter 2.2, according to which ASR in translation should not be understood only as a tool for first-draft dictation, but also as a possible aid for revision and navigation in hybrid CAT-tool workflows combining speech, mouse and keyboard input, where this seemed more convenient.

The purpose of this step was not to generate a separate body of methodological data, but to familiarize participants more directly with the demonstrated tools before they completed the questionnaire. In this sense, the try-out served as an intermediate step between the presentation and the questionnaire, allowing participants to relate the demonstrated functions more concretely to their own professional practice.

### **3.5 Questionnaire (EP RO Unit, DG TRAD)**

Following the presentation and the optional try-out, participants were asked to complete a questionnaire on ASR. The questionnaire was designed to capture perceptions related to the thesis framework of relevance, usability and acceptance, with a focus on Romanian-language institutional workflows.

The questionnaire included:

- Likert-scale questions on understanding of ASR, perceived relevance for professional and private tasks, perceived usability of the demonstrated solutions and willingness to try an institutionally approved ASR solution;
- Multiple-choice questions on previous ASR experience, which solutions seemed most relevant, which tasks ASR could realistically support and perceived barriers to adoption in the institutional context;
- Open questions allowing participants to describe perceived advantages and disadvantages, practical concerns, and views on future integration in the EP translation workflow.

To reflect realistic Romanian-language conditions, the presentation, optional try-out and questionnaire involved translators with different regional accents and one Lithuanian non-native Romanian-speaking professional. Since the theoretical foundation highlights pronunciation variability as a factor that can influence recognition quality and correction effort, including different accent profiles made it possible to capture whether participants expect accent, pauses, or pronunciation to affect ASR performance in their own case.

The questionnaire results are presented in subchapter 4.4, and Appendix D contains the questionnaire distributed to the participants, as well as the compiled results in two versions: the Romanian original and an English translation.

### **3.6 Methodology summary**

Taken together, the methodology combines two interviews, a presentation, an optional follow-up try-out and a questionnaire in order to examine the three forms of ASR use identified above. The two interviews provide complementary evidence on existing institutional ASR solutions, namely an expert perspective on live ASR and a user perspective on asynchronous transcription work. The presentation, try-out and questionnaire extend the analysis to active dictation and voice-command use in a CAT-tool context and make it possible to move from demonstration to direct engagement and then to structured participant feedback.

By combining these empirical steps, the methodology allows the findings to be interpreted consistently within the thesis framework of relevance, usability and acceptance, while also identifying concrete barriers and possible good practices for future integration of ASR in Romanian-language workflows.



## 4. Findings

The following chapter presents the findings in the same order as the empirical steps outlined in chapter 3. It begins with the two interviews, which address live / synchronous ASR (the interview with the Head of the EP's Speech-to-Text Unit) and asynchronous ASR (the complementary interview on Verba). It then turns to the presentation and to a short follow-up try-out, which both served as intermediate steps before the questionnaire on active ASR use. Because the try-out functioned primarily as preparatory hands-on exposure rather than as a fully independent methodological step, it is treated only briefly in subchapter 4.4. Throughout the chapter, the findings are interpreted through the analytical framework of relevance, usability and acceptance developed in chapter 2. The underlying interview transcripts, presentation materials and questionnaire documentation are provided in Appendices A-D.

### 4.1 Findings from the interview with the Head of the EP STT Unit

During the interview with Mr. Dragoş-Lucian Ciobanu, Head of the Speech-to-Text Unit, several relevant aspects emerged regarding how ASR is currently implemented, developed and evaluated within this unit of the European Parliament. In line with the methodological component addressing RQ1, the interview provided an expert perspective on current institutional ASR solutions in the European Parliament, going beyond the limited information available in institutional reporting. The full documentation of this interview is reproduced in Appendix A.

First, at a general level, the interviewee described the Speech-to-Text Unit's main task as the development, implementation and provision of ASR applications that ensure practical, instantaneous live transcription and translation in all 24 official languages of the European Union. More specifically, the interview made clear that, within this institutional context, live ASR is not framed as a replacement for interpreters. Instead, its main purpose is accessibility, especially for people with disabilities. In this respect, the interviewee referred primarily to people with hearing impairments, while also mentioning other conditions that may make it difficult to follow parliamentary events in real time, such as ADHD. According to the interviewee, around 52 million citizens of the European Union are affected by disabilities that influence hearing and participation in the democratic debate process. This makes accessibility a central institutional justification for live ASR and positions its relevance above all in relation to democratic access and participation, which are core principles of the European Union. This finding corresponds directly to the theoretical discussion in subchapter 2.3, where ASR in European institutions was described primarily as infrastructure for multilingual communication and accessibility. The interview therefore confirms that, in the EP, live ASR is first legitimized through its potential to widen access to parliamentary communication.

The interview also showed that the relevance of live ASR in the European Parliament is not limited to accessibility alone. The interviewee described additional practical advantages for Members of the European Parliament, staff, secretariats and accredited assistants. These internal users may need to follow more than one parliamentary event at the same time or recover immediate information they missed during a meeting because of distraction or noise. The interviewee gave examples of Members following another committee meeting remotely while



physically attending a different one or checking parliamentary discussions on their phones while travelling without switching on the sound. Mr. Ciobanu also gave a personal example of using the live transcript to recover part of a German intervention about Romania that he had not followed from the beginning, by scrolling back through the live transcription and consulting the translated output in English. These examples show that live ASR in the EP is not only a tool for disability-related accessibility, but also a support tool in a multilingual institution where many events take place simultaneously, in different languages, and where written and translated access to spoken content can be useful in several situations. This also supports the methodological distinction made in chapter 3 between live / synchronous ASR, active ASR use and asynchronous ASR. The interview clearly positions the Speech-to-Text Unit within the first category and relates its relevance to this specific category of ASR.

The interview then moved from the general role and advantages of live ASR to the concrete way in which this technology is currently implemented within the unit. The interviewee explained that the Speech-to-Text Unit has developed and implemented its current institutional ASR solution: a live Speech-to-Text platform based on LinTO, an open-source, self-hostable voice AI and transcription platform, described as being used by the European Commission and the European Parliament (LINAGORA AI n.d.)<sup>10</sup>. This platform is also combined with separate modules for language identification, transcription and translation. For the translation component, the unit currently uses Microsoft Azure AI Translator, a cloud-based neural machine translation service. According to Microsoft's official documentation, this service supports the translation of text and documents through API-based access and also offers possibilities for customised neural translation through Custom Translator and terminology-oriented dictionary features (Microsoft n.d.)<sup>11</sup>. Further, Microsoft states that customer data submitted for translation is not stored permanently and that organisational data remains within the customer's Azure subscription (Microsoft 2025)<sup>12</sup>.

This finding adds an important nuance to the theoretical discussion in subchapters 2.2 and 2.3, where cloud-based commercial solutions were presented as potentially problematic in institutional environments because they may raise concerns about confidentiality, data protection and control over speech or text data. The present interview does not invalidate that concern but suggests that the distinction is not simply between institutional and commercial tools, or between local and cloud-based solutions. Rather, the key issue is whether a commercial component can be integrated into a controlled institutional environment, with clear rules on data handling, access and domain adaptation.

In this sense, the current Speech-to-Text solution in the EP appears as a modular system that combines institutional infrastructure with a commercial translation component that is more governed and configurable than an unregulated public web service. This gives more precise information than the general institutional reporting discussed in chapter 2.3 and suggests that the current EP solution does not rely on a fully internal end-to-end system, but a modular setup shaped by both technical and data governance considerations. The interviewee further explained that the idea of building a fully internal ASR solution without any commercial

---

<sup>10</sup> <https://www.linagora.ai/en/solutions/linto>

<sup>11</sup> <https://learn.microsoft.com/en-us/azure/ai-services/translator/>

<sup>12</sup> <https://learn.microsoft.com/en-us/azure/ai-services/translator/faq>

component has itself become difficult, because such a system would be more expensive and has already led to conflicting discussions with the IT side of the EP.

Further, the interview immediately added an important clarification. Although a commercial resource is currently used, the interviewee outlined that such a resource could not be used “off the shelf” in the European Parliament. Mr. Ciobanu gave three reasons for this. First, speech and language data gathered broadly from the internet, often through data mining and used to train or improve AI-based ASR systems, do not match the type of speech and language data needed for parliamentary events. In the case of Romanian, he pointed out that much publicly available online language is strongly influenced by informal and entertainment-oriented usage and therefore differs from parliamentary discourse. Second, the grammar and actual patterns of language use in Parliament differ from what such systems encounter in broad internet data. Third, for several languages there are important differences between the varieties spoken within and outside the European Union, while AI is not yet precise enough to distinguish between them reliably. He gave the examples of French, Portuguese and Spanish, where European and non-European varieties may differ considerably. The interviewee also linked this issue to data protection, suggesting that some ASR providers may rely on data collected outside the European Union, where the rules are less strict and data may be obtained more extensively through data mining.

These remarks correspond closely to the theoretical discussion in chapters 2.1 and 2.3. In chapter 2.1, this thesis argued that large-vocabulary ASR is not the same as domain-appropriate ASR and that usefulness depends on whether the language model can be adapted to the relevant context. In chapter 2.3, it was further argued that institutional ASR depends on in-domain parliamentary speech and aligned transcripts rather than on general-language coverage alone. The interview confirms these points in a very concrete way and suggests that, in the EP, usability is shaped not simply by the existence of multilingual ASR, but by how well the system matches parliamentary discourse, institutional phraseology and institutional linguistic norms.

The issue of domain-specific language adaptation was addressed more directly in the next part of the interview. When asked how the commercial resources had been adapted from an institutional terminological point of view, the interviewee explained that this had been done through customisation. Mr. Ciobanu outlined that parliamentary discourse contains many acronyms, proper names and specific terms that do not normally appear in other contexts and therefore need to be explicitly introduced and trained into the system.

The examples of ENVI and PETI are especially relevant here. In ordinary English, these forms may be associated with other meanings (envy and petty), but in the parliamentary context they refer to committees and institutional entities. This finding corresponds to the theoretical discussion in chapter 2.1, where large-vocabulary ASR was distinguished from domain-appropriate ASR. The interview confirms that, even when a general ASR resource performs well in broad terms, its usefulness in an institutional context still depends on whether it can be adapted to institutional terminology and phraseology.

The interviewee also made clear that this adaptation process is still ongoing and that the unit is currently customizing and training AI-based language models for their ASR solution. This shows that the live Speech-to-Text platform should not be understood as a finished product in a strict sense, but rather as an institutional service that will be further improved.

The interview then moved to the broader question of challenges and limitations. The interviewee explained that the development and improvement of the Speech-to-Text platform is difficult for several reasons: human, technical, budgetary and institutional. From the human-resources point of view, it is not easy to involve professional linguists from other units or Directorates in the customization process, because DG TRAD staff are already overloaded with their regular work. For this reason, the unit relied mainly on trainees, while Parliament linguists were involved only for higher-level or more delicate tasks. From the technical point of view, the interviewee mentioned error management, ensuring good sound quality, finding and customizing suitable language models and building sufficiently large databases, especially for languages with fewer available resources. From the institutional point of view, these technical and linguistic tasks are made more difficult by limited staff, limited funding, coordination needs across services and compliance requirements. In this sense, the interview shows that developing an institutional ASR solution is not only a technical task. It is also a resource-intensive and institutionally complex process. Even so, the interview also suggests that the unit is trying to improve the platform step by step under these constraints.

The next part of the interview focused further on Romanian-specific aspects. When asked whether Romanian should be considered a low-resource language for language-data collection, the interviewee rejected this classification and stated that Romanian had performed “surprisingly well” in the tests carried out on speeches by Romanian Members of the European Parliament. This statement nuances the broader picture presented in the theoretical subchapter 2.4, where Romanian was described as historically affected by resource limitations in ASR research. The interview suggests that Romanian may currently perform better than expected.

Mr. Ciobanu also added that performance could depend on accent. However, when asked whether, for example, a strong Moldavian accent could create difficulties in recognising spoken Romanian, he argued that such difficulties are less relevant in this institutional context. The Romanian parliamentary setting is shaped by a cultural and institutional norm of speaking in a standardized and careful way. According to him, Romanian MEPs generally do not have an incentive to speak at the rostrum in a strong local dialect. On the contrary, official parliamentary situations favour a form of Romanian that is as standard and correct as possible. He contrasted this with other contexts, such as the Swiss Parliament, where stronger local varieties may be more present in formal speech. This observation corresponds closely to the discussion in chapter 2.1 on speaker variability. It adds another layer and suggests that, in the Romanian parliamentary context, some of the variability that often creates difficulty for ASR may be reduced by the institutional speech norm itself. In this sense, Romanian seems to benefit not only from technical support, but also from the relative standardization of institutionally acceptable parliamentary speech.

Further, although Romanian was described as performing well overall, the interviewee stated that the unit had not had the necessary resources to dedicate very detailed attention to Romanian. The Speech-to-Text Unit is small, and much of its effort has gone into getting the system up and running. Romanian has therefore not yet been customised in depth, and its errors have not yet been systematically recorded and analysed. Even so, the interviewee explained that a general error-rate check placed Romanian among the languages with the lowest error rates and that both the internally used parliamentary Romanian data and the Microsoft Azure data were “quite good”. These statements refine the Romanian-focused discussion in subchapter 2.4.

Romanian appears promising in this specific institutional setting, but the evidence remains more general than detailed. The interview therefore partially reduces the Romanian-specific evidence gap but does not eliminate it completely in the absence of a more detailed evaluation of its performance.

This unevenness becomes even clearer in the next exchange on the evaluations carried out by voluntary trainees due to a lack of staff. The interviewee explained that 53 trainees from all 24 languages were involved, but that their contributions differed enormously, with some completing all assigned tasks and others doing none. In the case of Romanian, the contribution was relatively limited because the Romanian translation unit had only one trainee instead of the two originally expected. This shows that support across languages is uneven not only in theory, but also in very concrete operational terms. The development of multilingual institutional ASR depends not only on technical infrastructure and speech data, but also on the availability of people who can contribute to adaptation and checking.

A further key issue addressed in the interview is evaluation. When asked whether any experiments had been carried out to compare a fully automated live-ASR workflow with a workflow involving ASR plus active human correction, such as a respeaker, the interviewee answered that no such benchmarking had yet been conducted because of lack of resources. He explained that the Speech-to-Text Unit consists of only three people on the technical side, one administrative person, one communication person and the Head of Unit himself, and that all are overloaded. This finding is especially relevant in relation to chapter 2.5, where limited transparency on how institutional ASR is evaluated in practice was identified as a research gap. The interview shows that the lack of benchmarking is not simply an omission in reporting, but is also linked to very concrete operational limits.

In the same context, the interviewee also outlined an internal prioritisation of languages. English received the highest level of attention because many of the events covered by the unit were predominantly in English. French and German also received focused attention for event-related reasons. At the other end of the spectrum, Maltese and Irish were described as especially difficult cases because commercial providers offer almost no support for them. Romanian, together with languages such as Italian, Spanish and Polish, was placed in a middle category. This helps contextualise the earlier Romanian findings: Romanian appears to perform relatively well in the present setting even though it has not been among the most heavily prioritised languages for customisation. This suggests once more that parliamentary Romanian may currently be comparatively favourable for live ASR, but also that its positive performance should not be interpreted as the result of very detailed attention to language-specific optimisation and evaluation.

Another practical point clarified by the interview concerns the use of the application itself. When asked whether the application can be accessed publicly at any time, the interviewee explained that it is designed specifically for live use. It is intended for participants following debates during the event itself and cannot be accessed afterwards. This operational detail is important because it reinforces the distinction made in chapter 3 between live / synchronous ASR and asynchronous ASR. The Speech-to-Text platform is designed to provide immediate access during the event, not long-term post-event documentation. This live-only orientation is directly tied to its accessibility purpose, but it also creates difficulties for later checking and systematic evaluation. He explained that, when the unit wants to test and verify at a later stage

the quality of the transcriptions generated by their live ASR platform, data-protection issues arise immediately and require authorization. For some non-public events, the unit did not even have permission to save the transcript afterwards in order to perform a proper quality check. In such cases, verification could be carried out only live within the unit itself, each colleague following their own working languages live. This finding addresses the theoretical discussion presented in chapters 2.2 and 2.3, where confidentiality, data protection and access control were presented as key institutional constraints. The interview confirms that these concerns are not abstract. They have direct consequences for how evaluation and improvement can be carried out.

The interviewee also mentioned a concrete attempt to improve evaluation through a public ECON meeting that had been broadcast live and therefore did not raise the same data-protection difficulties in accessing at a later stage the concrete ASR output: the transcripts. In this case, the unit intends in the future to send the material to all 24 linguistic units so that they could check word error rate, latency and general performance. However, he made clear that this requires considerable human time and is difficult to organize institutionally and logistically. Multilingual evaluation in the EP depends on the institution's ability to mobilize coordinated human resources across all 24 languages. It therefore strengthens the methodological rationale of the present thesis: expert interviewing is necessary precisely because institutional documentation does not show this full operational complexity behind such institutional services.

Further, according to Mr. Ciobanu, the European Parliament is a difficult environment for deploying live ASR also because Members are very risk-averse when it comes to the possibility that instantaneous transcription and translation containing errors could misrepresent their speeches. In this respect, the interview highlighted a central contradiction of live ASR in the EP. In order to fulfil its accessibility function, the system must deliver automatic speech recognition - transcription and translation almost immediately and without prior human intervention. According to the interviewee, users depending on such support expect the spoken content to be identified, transcribed and translated within a maximum of three seconds. At the same time, however, the institutional and political context makes inaccurate output particularly sensitive. The same feature that gives live ASR its value: immediate delivery also makes it risky, because the output reaches users before it can be checked or corrected by a human.

The interviewee illustrated the consequences of this tension through concrete examples. He explained that, in politically sensitive contexts, even a small omission may completely change the meaning of a statement. If, for example, an MEP says that a certain action is “not politically accepted” and the system misses the negation, the output may suggest the opposite political position. He also pointed to the risks created by language switching. In multilingual parliamentary speech, speakers may move from one language to another and the system may become confused. Mr. Ciobanu gave the example that, if a speaker uses the English word library while otherwise speaking Romanian, the ASR application may produce librărie (bookshop) instead of bibliotecă (library), because of the phonetic similarity. These examples show that, in the institutional context of the EP, acceptance does not depend only on average quality or general usability. It also depends on the possible consequences of individual errors, especially when political meaning, public interpretation and the speaker's image are at stake.

This point also helps clarify why live ASR should not be evaluated according to exactly the same logic as other types of ASR use discussed in this thesis. In the case of active ASR use, such as dictation or voice commands in a CAT-tool or office environment, users may tolerate a certain level of error because the output remains under their direct control and can be revised before further use. Likewise, in asynchronous ASR, such as post-event transcription workflows, there is time for checking, editing and correction. In live ASR, by contrast, the usefulness of the system depends precisely on its ability to provide immediate access without human mediation. At the same time, this is exactly what makes it risky in the EP context: when a machine produces the output on its own and in real time, errors can appear before anyone has the chance to correct them. The interview therefore supports the methodological distinction made in chapter 3 between live / synchronous ASR, active ASR use and asynchronous ASR, because these different forms of use involve different expectations, different risks and different acceptance conditions.

This distinction became even clearer in the final exchange of the interview, when the discussion turned to Verba. As already outlined in chapter 2.3, Verba is another ASR solution used in the European Parliament, in the transcription workflow for plenary speeches. In other words, Verba belongs to a different ASR use case from the live Speech-to-Text platform developed by the Speech-to-Text Unit. While the latter is designed mainly for live accessibility and multilingual support during parliamentary events, Verba is used in a more controlled transcription context, where human correction is possible. The interview explicitly pointed out that, for plenary sessions, Verba is already used in an active way, meaning that someone is present to check and correct the output so that the final text does not contain something the Member did not intend to say. Mr. Ciobanu explained that this is precisely why, at plenary level, the live Speech-to-Text application is not preferred: the political and reputational risk of immediate, uncorrected output is considered too high.

At the same time, the interviewee also outlined that Verba cannot fulfil the same function as the live Speech-to-Text platform. Verba is not intended for the wider public and it is not automatically translated into all 24 official languages. For this reason, it cannot provide the same form of immediate multilingual accessibility for persons with disabilities. This comparison makes the distinction between live / synchronous ASR and asynchronous ASR even more concrete. In the case of plenary debates, where political sensitivity is particularly high, a more controlled workflow with human correction appears more acceptable. In the case of the Speech-to-Text platform, however, the value of the system lies in providing immediate access during events, including events that are not necessarily as politically exposed as plenary speeches and especially are not publicly accessible afterwards. The interview therefore shows that these two institutional ASR solutions should not be understood as competing versions of the same tool, but as different solutions designed for different institutional purposes, with different expectations regarding relevance, usability and acceptance.

Overall, the interview helps cover several of the gaps identified in the theoretical foundation. It shows more concretely that the EP Speech-to-Text platform is conceived first of all as a live accessibility service and only secondarily as a practical support tool for Members and staff. It also clarifies that the platform is built as a multimodal institutional solution, combining LinTO-based infrastructure with separate recognition, transcription and translation components, while also relying on a controlled commercial translation resource. Another

important finding is that commercial ASR resources cannot simply be used as they are, but require adaptation to parliamentary language, terminology and institutional conditions. The interview further brings a first Romanian-focused perspective from within the EP by suggesting that Romanian performs relatively well in this setting, partly because parliamentary Romanian tends to be more standardized, while at the same time showing that Romanian has not yet been examined in detail through systematic language-specific evaluation. In addition, the interview makes visible several practical constraints that are largely absent from institutional reporting, such as limited staffing, dependence on trainees, uneven support across languages, restricted possibilities for later quality checks and the difficulty of benchmarking live ASR across all 24 languages. Finally, it clarifies more sharply the difference between live and asynchronous ASR in the European Parliament: live ASR is valuable because it gives immediate access without human intervention, but this same immediacy also makes it more sensitive to error and therefore more difficult to accept in politically exposed contexts. Taken together, these findings respond directly to RQ1 and specify the live, accessibility-oriented side of the institutional ASR landscape outlined in subchapter 2.3.

At the same time, this part of the empirical material also has clear limitations. It reflects primarily the perspective of the Head of the Speech-to-Text Unit and does not provide direct evidence from end users of the live platform, such as accessibility users or participants in parliamentary events. It also does not include detailed quantitative evaluation data, a systematic Romanian-specific error typology or a benchmark comparison with human-mediated live workflows.

Nevertheless, it makes a substantial contribution to the present thesis because it shows more clearly how live institutional ASR is shaped in practice in the European Parliament and why its relevance, usability and acceptance must be understood in relation to its specific live and accessibility-oriented function.

## **4.2 Findings from the Verba interview (EP RO Unit, DG TRAD)**

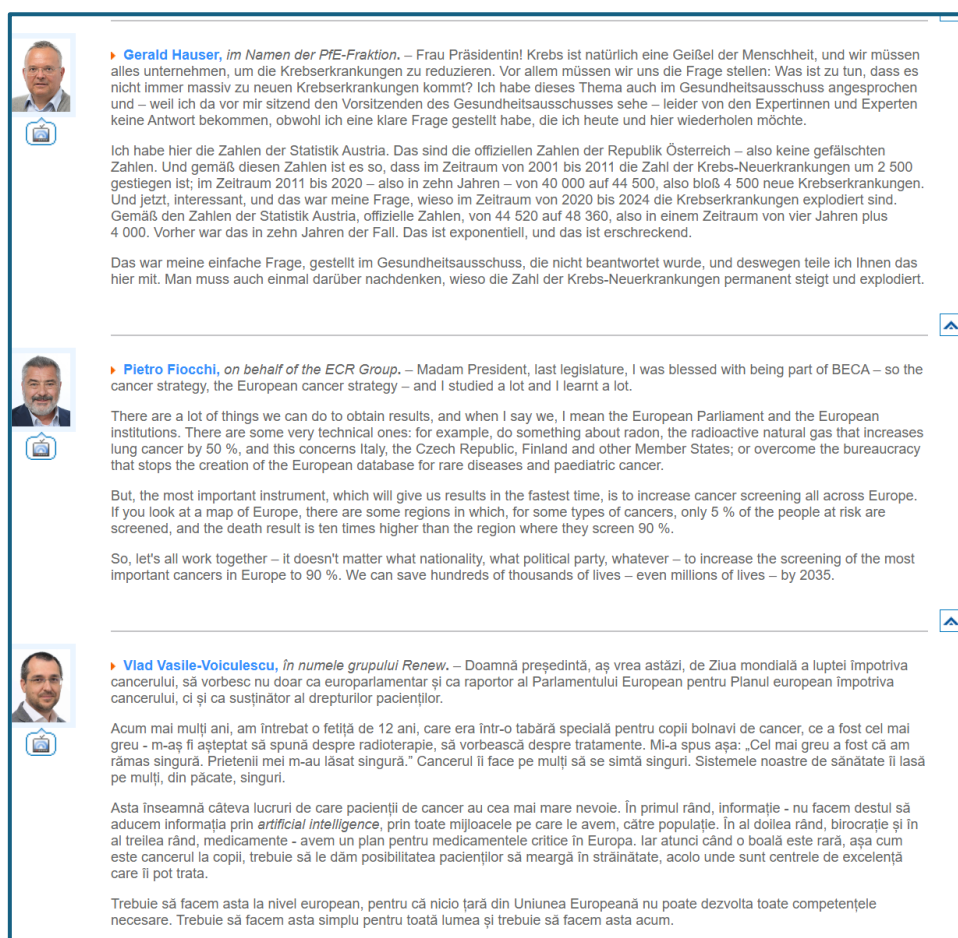
The interview with the anonymous experienced Romanian proofreader from the European Parliament Romanian Translation Unit (DG TRAD) provided a complementary perspective on institutional ASR in the European Parliament. Unlike the previous interview, which focused on the institutional development and evaluation of live ASR as an accessibility-oriented service, this interview brought in the perspective of a daily user working with Verba in a concrete production setting. It also concerned a different institutional use case, namely asynchronous ASR for strictly internal transcription and documentation workflows.

In line with the methodological rationale outlined in chapter 3.2, the interview contributes to RQ1 by adding a user-based perspective on another current institutional ASR solution in the European Parliament and to RQ3 by providing a Romanian-specific perspective from a professional user working in the EP context. The full interview documentation is provided in Appendix B.

A first important point concerns the function of Verba itself. The interviewee described it as an application used to transcribe speeches by Members of the European Parliament and, in some cases, by other high-level speakers, such as Commissioners. The tool is used mainly during

plenary sessions, but also in some committee-related or similarly important settings. According to the interviewee, in the Romanian Unit proofreaders work within the Verba application on these automatically generated transcriptions in order to prepare them for official publication. The revised texts are then published as part of the verbatim reports of proceedings, often referred to by the French abbreviation CRE (*compte rendu in extenso*), on the European Parliament's plenary website, where they are searchable by date, speaker and keyword and are also linked to the corresponding video segments of the debate (European Parliament n.d.a)<sup>13</sup>.

Unlike the live Speech-to-Text platform, for which limited public interface screenshots are available, no comparable public representation of the Verba interface could be identified. In the case of Verba, the publicly accessible material concerns primarily the published output of the workflow rather than the internal tool itself.



The screenshot displays a list of three speakers and their interventions. Each entry includes a small profile picture, a speaker icon, the speaker's name and affiliation, and the text of their speech in the original language.

- Gerald Hauser, im Namen der PFE-Fraktion.** – Frau Präsidentin! Krebs ist natürlich eine Geißel der Menschheit, und wir müssen alles unternehmen, um die Krebserkrankungen zu reduzieren. Vor allem müssen wir uns die Frage stellen: Was ist zu tun, dass es nicht immer massiv zu neuen Krebserkrankungen kommt? Ich habe dieses Thema auch im Gesundheitsausschuss angesprochen und – weil ich da vor mir sitzend den Vorsitzenden des Gesundheitsausschusses sehe – leider von den Expertinnen und Experten keine Antwort bekommen, obwohl ich eine klare Frage gestellt habe, die ich heute und hier wiederholen möchte.  
Ich habe hier die Zahlen der Statistik Austria. Das sind die offiziellen Zahlen der Republik Österreich – also keine gefälschten Zahlen. Und gemäß diesen Zahlen ist es so, dass im Zeitraum von 2001 bis 2011 die Zahl der Krebs-Neuerkrankungen um 2 500 gestiegen ist; im Zeitraum 2011 bis 2020 – also in zehn Jahren – von 40 000 auf 44 500, also bloß 4 500 neue Krebserkrankungen. Und jetzt, interessant, und das war meine Frage, wieso im Zeitraum von 2020 bis 2024 die Krebserkrankungen explodiert sind. Gemäß den Zahlen der Statistik Austria, offizielle Zahlen, von 44 520 auf 48 360, also in einem Zeitraum von vier Jahren plus 4 000. Vorher war das in zehn Jahren der Fall. Das ist exponentiell, und das ist erschreckend.  
Das war meine einfache Frage, gestellt im Gesundheitsausschuss, die nicht beantwortet wurde, und deswegen teile ich Ihnen das hier mit. Man muss auch einmal darüber nachdenken, wieso die Zahl der Krebs-Neuerkrankungen permanent steigt und explodiert.
- Pietro Fiocchi, on behalf of the ECR Group.** – Madam President, last legislature, I was blessed with being part of BECA – so the cancer strategy, the European cancer strategy – and I studied a lot and I learnt a lot.  
There are a lot of things we can do to obtain results, and when I say we, I mean the European Parliament and the European institutions. There are some very technical ones: for example, do something about radon, the radioactive natural gas that increases lung cancer by 50 %, and this concerns Italy, the Czech Republic, Finland and other Member States; or overcome the bureaucracy that stops the creation of the European database for rare diseases and paediatric cancer.  
But, the most important instrument, which will give us results in the fastest time, is to increase cancer screening all across Europe. If you look at a map of Europe, there are some regions in which, for some types of cancers, only 5 % of the people at risk are screened, and the death result is ten times higher than the region where they screen 90 %.  
So, let's all work together – it doesn't matter what nationality, what political party, whatever – to increase the screening of the most important cancers in Europe to 90 %. We can save hundreds of thousands of lives – even millions of lives – by 2035.
- Vlad Vastile-Voiculescu, în numele grupului Renew.** – Doamnă președintă, aş vrea astăzi, de Ziua mondială a luptei împotriva cancerului, să vorbesc nu doar ca europarlamentar şi ca raportor al Parlamentului European pentru Planul european împotriva cancerului, ci şi ca susţinător al drepturilor pacienţilor.  
Acum mai mulţi ani, am întrebat o fetiţă de 12 ani, care era într-o tabără specială pentru copii bolnavi de cancer, ce a fost cel mai greu - m-aş fi aşteptat să spună despre radioterapie, să vorbească despre tratamente. Mi-a spus aşa: „Cel mai greu a fost că am rămas singură. Prietenii mei m-au lăsat singură.” Cancerul îi face pe mulţi să se simtă singuri. Sistemele noastre de sănătate îi lasă pe mulţi, din păcate, singuri.  
Asta înseamnă câteva lucruri de care pacienţii de cancer au cea mai mare nevoie. În primul rând, informaţie - nu facem destul să aducem informaţia prin *artificial intelligence*, prin toate mijloacele pe care le avem, către populaţie. În al doilea rând, birocrăţie şi în al treilea rând, medicamente - avem un plan pentru medicamentele critice în Europa. Iar atunci când o boală este rară, aşa cum este cancerul la copii, trebuie să le dăm posibilitatea pacienţilor să meargă în străinătate, acolo unde sunt centrele de excelenţă care îi pot trata.  
Trebuie să facem asta la nivel european, pentru că nicio ţară din Uniunea Europeană nu poate dezvolta toate competenţele necesare. Trebuie să facem asta simplu pentru toată lumea şi trebuie să facem asta acum.

Figure 7: Excerpt from the verbatim report of proceedings, 9 February 2026, Strasbourg (European Parliament 2026a)

Figure 7 shows the European Parliament's verbatim report interface (2026a)<sup>14</sup>, where individual plenary interventions are displayed in separate speaker-based entries, each accompanied by a photo, speaker identification and the speech text in the original language.

<sup>13</sup> <https://www.europarl.europa.eu/plenary/en/debates-video.html>

<sup>14</sup> [https://www.europarl.europa.eu/doceo/document/CRE-10-2026-02-09-ITM-015\\_EN.html](https://www.europarl.europa.eu/doceo/document/CRE-10-2026-02-09-ITM-015_EN.html)



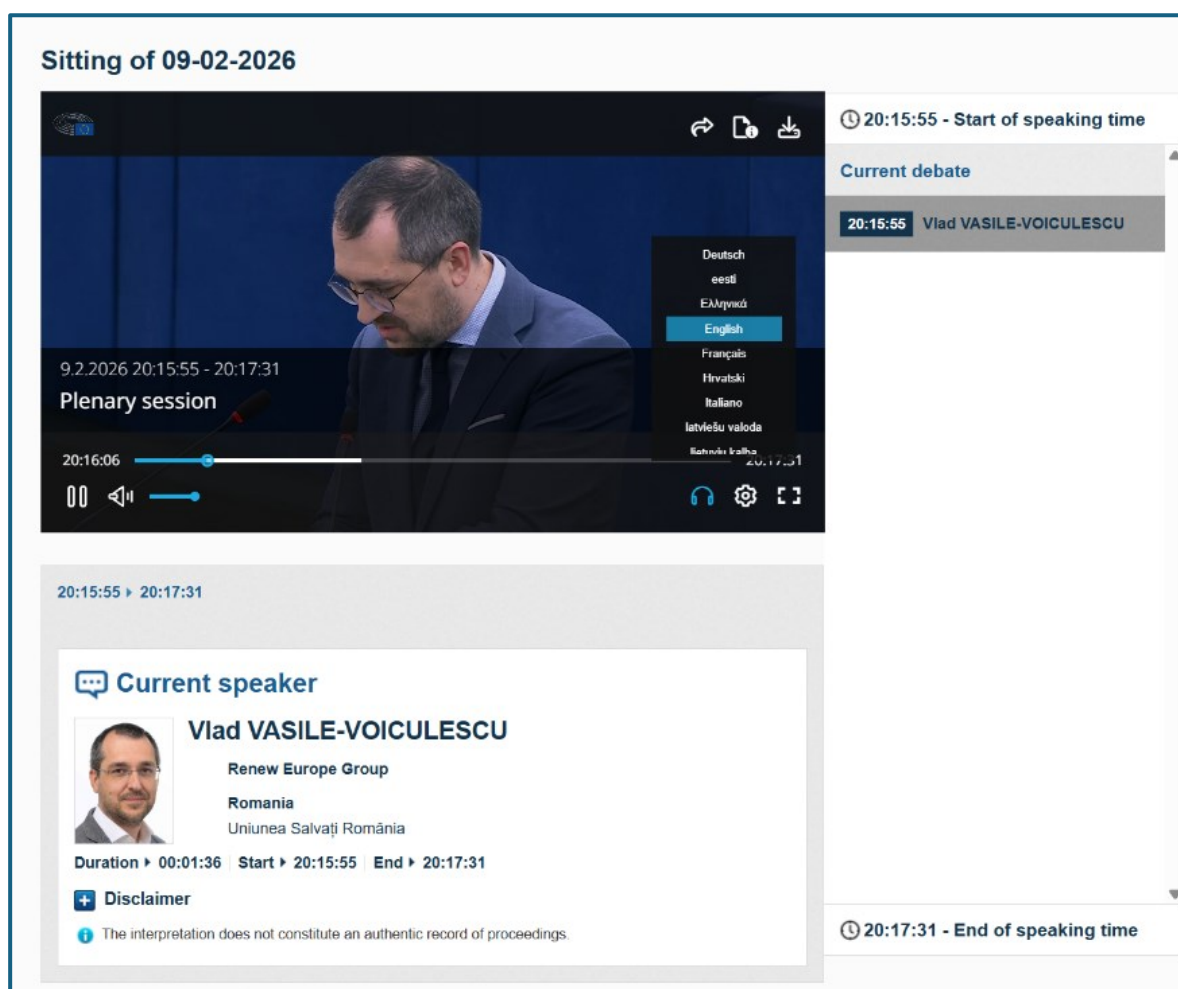


Figure 8: Video-on-demand entry for Vlad VASILE-VOICULESCU, Sitting of 9 February 2026 (European Parliament 2026b)

Figure 8 shows the corresponding video-on-demand interface (2026b)<sup>15</sup>, where the same intervention can be consulted together with speaker metadata and the precise timing of the speech.

This is a key finding in relation to the distinction drawn in chapter 3 between live / synchronous ASR, asynchronous ASR and active ASR use. Verba clearly belongs to the second category. The proofreader does not use speech to produce new text, as in dictation, and does not rely on instant public output, as in live ASR. Rather, the proofreader works on a written output that has already been generated and that remains subject to human correction before publication. At the same time, the later publication of these revised transcriptions as searchable verbatim reports linked to the corresponding video segments shows that Verba is embedded in a broader institutional documentation workflow. This confirms that institutional ASR in the EP cannot be treated as a single use case. Different tools serve different functions and their relevance must be judged according to different expectations regarding relevance, usability and acceptance.

15

<https://www.europarl.europa.eu/plenary/en/vod.html?mode=unit&vodLanguage=EN&playerStartTime=20260209-20:15:55&playerEndTime=20260209-20:17:31#>

The interviewee's comments made further clear the relevance of this ASR solution. The interviewee recalled a situation in which the application was not functioning and the speeches transcripts were not available, which meant that the speeches had to be written down manually. This was described as "awful". That reaction is highly revealing, suggesting that Verba is no longer perceived as an optional convenience, but as a necessarily integral support in the workflow. Once the user is used to working with automatically generated text that only needs revision, returning to manual transcription feels exhausting and inefficient. In terms of relevance, this is a strong indicator that Verba has become meaningfully embedded in the proofreader's everyday practice.

Further, the interviewee stated very clearly that Verba makes the transcription process faster and more efficient and also explicitly compared Verba with CRE, the previous tool used before Verba and stated that Verba is "much better". As discussed in chapter 2.3, public documentation confirms that Verba replaced CRE and that end-to-end tests were positive, but it does not show how users experience this change in everyday work. The present interview therefore helps fill part of that gap by showing that, from the perspective of an experienced Romanian proofreader, the newer system is not experienced just as a technical replacement, but as a more useful and more efficient working tool.

The user also linked this gain in efficiency to the quality of the output itself. According to the user, Verba recognizes spoken Romanian better than the previous CRE and seems to perform better even when speeches are not ideally pronounced or well formulated. This observation is relevant in relation to the theoretical discussion in chapter 2.1, where continuous speech recognition was described as difficult precisely because real speech includes pauses, reformulations, variable speed and unclear boundaries between sentences. The user's account suggests that Verba is perceived as more efficient than its predecessor under such realistic speaking conditions.

A further interesting point emerged when the interviewee described the new Verba, in contrast with the previous CRE, as a system in which one can already "see a bit of the hand of AI" and that appears to "learn". According to the user, the application seems to improve over time and sometimes produces solutions similar to those a human proofreader would make. The interviewee mentioned, for example, that Verba may remove certain repeated sounds or hesitation-like forms, place punctuation "quite well", and in some cases even correct grammatical agreement, so that the output appears partly normalized toward a more publishable written form.

This corresponds to the theoretical discussion in chapter 2.1 on large-vocabulary ASR, language modelling and the role of AI in the more recent development of speech recognition. As outlined in the theoretical foundation, AI-based methods have made ASR systems more adaptable and better able to cope with realistic speech conditions. In this sense, the interviewee's impression that Verba "learns" may reflect improved language modelling, domain adaptation, terminology updates or post-processing. The finding is therefore that Verba is perceived as reducing part of the proofreader's effort to correct transcripts by delivering output that is already partially processed and normalized. At the same time, this perception must be interpreted carefully. It does not provide direct technical evidence that the system is self-learning, since the interviewee also stated that the internal functioning of the application is not known to them and that only the PV-CRE Unit would know whether specific terms are

introduced manually or how the system is updated. The existence of the PV-CRE Unit itself is documented in a European Parliament annual activity report, which refers explicitly to the “Unité ‘Procès verbaux et comptes rendus de la plénière’” and to its cooperation with DG TRAD (European Parliament 2010: 24-27).

The next part of the interview clarified more precisely how the interviewee evaluated Verba’s accuracy in the specific context of plenary sittings. The interviewee described the application as “fairly accurate” overall, but also pointed to a clear limitation related to parliamentary terminology. According to the user, Verba does not always recognize all names of EU programmes and other newer or less recurrent institutional designations. Some more recurrent expressions, such as Green Deal, were said to be recognized, whereas others are not always identified correctly. The interviewee also suggested that names of committees, Members and Commissioners had probably been introduced from the start by the PV-CRE Unit, while other denominations remain more problematic. This corresponds to the theoretical discussion in chapters 2.1 and 2.2, where it was argued that broad vocabulary coverage is not the same as domain-appropriate performance and that usefulness in institutional settings depends on how well the system handles specialized terminology, acronyms and proper names. The finding is therefore that Verba is perceived as fairly accurate in plenary work overall, but that one of its clearest limits remains the recognition of newer or less established institutional names.

The interview then moved to the question of typical mistakes in Romanian. Here, the interviewee did not identify a clear category of specifically Romanian errors, but linked most remaining problems to speaker behaviour rather than to the language itself. According to the user, if the speaker repeats something, the automatic transcription may also repeat it, although in some cases Verba appears to remove obvious repetitions or stuttering. Likewise, some word confusions still occur, but these were attributed mainly to pronunciation, in the sense that the application “understands something else”. This corresponds to the theoretical discussion in chapter 2.1 on realistic speech conditions, disfluency and speaker variability. Parliamentary speech is not always linear or ideally delivered, and even in a relatively formal setting ASR remains sensitive to pauses, repetition and pronunciation. The finding is therefore that the interviewee does not perceive Romanian-specific error patterns as the main problem; rather, the quality of the output depends above all on how clearly and carefully the speech is delivered.

This point becomes even clearer in the following part of the interview, where the interviewee described the amount of correction needed in practice. According to the user, there are speeches in which very little needs to be corrected. In general, when speakers read, have correct pronunciation and formulate their sentences properly, almost no intervention is required apart from perhaps a comma or another very minor correction. By contrast, more intervention is needed when speakers speak more freely, fail to complete one sentence before beginning another, or pronounce words less clearly. The interviewee also emphasised that some pronunciation-related mistakes are not immediately obvious in the transcript and can only be detected by listening carefully to the speech. The need for human oversight remains evident. This corresponds directly to the methodological focus outlined in chapter 3.2, where post-correction effort was identified as a key indicator of usability. It also reflects the broader discussion in chapter 2.2, according to which ASR becomes professionally useful only when the effort needed for correction does not cancel out the efficiency gained. Taken together, these observations indicate that Verba can substantially reduce correction effort and support

efficiency in the transcription workflow, although not equally across all speeches. Its usability remains closely dependent on the way the intervention is delivered, especially in terms of clarity, sentence structure and pronunciation, so that human revision continues to be indispensable.

A particularly relevant Romanian-specific nuance emerged in the exchange on homophonous forms such as *sau* and *s-au*. In chapter 2.4, this thesis noted that Romanian ASR may encounter difficulties with orthographic ambiguity, hyphenation and clitic constructions, especially where different written forms correspond to the same pronunciation. Avram, Păiș and Tufiș (2020: 397–398) likewise explain that, in Romanian ASR, the raw output does not automatically account for hyphenation and that forms such as *s-au* may initially appear as *sau*, which means that the correct written form must be recovered on the basis of context through language modelling and post-processing. In the interview, however, the user stated that they had not noticed errors of this type and had not observed similar cases in which identical pronunciation led to an incorrect written form. This suggests that, at least from the perspective of this experienced proofreader, Verba does not rely only on acoustic recognition, but also benefits from contextual processing that can distinguish between plausible written alternatives.

The final exchange of the interview turned to possible improvements. The interviewee answered, partly ironically, that ideally the application should simply “recognise everything”, but then specified more concretely that new names and terms should either be learned by the system or introduced manually. This answer is consistent with the earlier remarks on parliamentary terminology and again points to lexical updating as one of the main areas where Verba could still improve. At the same time, the interviewee also made clear that they do not know how this process works internally and that only the PV-CRE Unit would know how specific terms are manually added. This confirms an important methodological limit of the present interview: it gives access to the user experience of Verba, but not to its technical architecture or to the exact institutional procedures behind its updating and customisation process.

Overall, the interview fills an important part of the gap identified in the theoretical foundation by providing a concrete Romanian-language user perspective on Verba as an institutional ASR solution. It shows more clearly that Verba should be understood as a controlled asynchronous transcription tool embedded in a broader parliamentary documentation and publication workflow, rather than as a public-facing live ASR service. From the perspective of an experienced Romanian proofreader, its main relevance lies in the fact that it has become an integral support in the plenary workflow, making transcription substantially faster and more manageable than manual work and clearly more efficient than the previous CRE tool. In terms of usability, the interview suggests that Verba performs fairly well overall in Romanian, including under realistic plenary conditions, and that it can already reduce part of the correction burden through better recognition, partial normalisation of the output and context-sensitive processing. At the same time, its performance remains uneven in relation to newer or less recurrent parliamentary terms, while the amount of post-correction still depends heavily on the speaker’s delivery, especially pronunciation, sentence structure and degree of spontaneity.

In this respect, the interview complements the findings of chapter 4.1 in an important way. Whereas the previous interview showed that live ASR in the EP is shaped primarily by accessibility goals and immediacy requirements, the Verba interview shows that in a more

controlled, asynchronous workflow the same institutional environment allows for greater acceptance because human revision remains possible before publication. Taken together, the two interviews therefore confirm more clearly that current institutional ASR solutions in the European Parliament should not be understood as competing versions of the same technology, but as different tools designed for different institutional purposes, with different expectations regarding relevance, usability and acceptance. They also complement each other in relation to Romanian: while chapter 4.1 suggested that Romanian performs relatively well in the EP's live ASR setting, the present interview adds a concrete user-based indication that Romanian output in Verba is generally manageable in practice, even though terminology and speaker-dependent variation still require attention.

At the same time, this part of the empirical material also has clear limitations. It is based on one short interview with one experienced Romanian proofreader and therefore cannot provide a broader picture across all Romanian users or across other language units. It reflects user perception rather than direct technical evaluation and does not include quantitative measures such as word error rate, time savings or a systematic Romanian-specific error typology. In addition, the interview gives insight into the practical experience of using Verba, but not into its internal architecture, training process or updating procedures, since the interviewee explicitly stated that such details were known only to the PV-CRE Unit. Nor does it include comparison with less experienced users or with users from other linguistic units, which would have made it possible to assess more systematically whether the positive experience reported here is typical or more individual.

Nevertheless, despite these limitations, the interview makes a substantial contribution to the present thesis because it adds a concrete practice-based account of how asynchronous institutional ASR is experienced in Romanian-language work and thereby complements the more development-oriented and live-ASR-focused perspective of subchapter 4.1.

### **4.3 Findings from the presentation for the EP RO Unit, DG TRAD**

The preparation of the presentation for the European Parliament Romanian Unit produced findings that go beyond the general publicly available descriptions of ASR tools and make more visible how ASR can, or cannot, fit institutional workflows in practice. In line with the methodological rationale outlined in subchapter 3.3, this empirical step addressed active ASR use for dictation and voice-enabled interaction. More specifically, it contributed primarily to RQ2 by exploring how ASR could fit everyday office tasks and CAT-tool work in a Romanian-language institutional setting. In this sense, the preparation of the presentation did not only serve to identify demonstrable ASR options but also revealed concrete institutional constraints that shape ASR relevance, usability and possible acceptance in the workplace. The full presentation documentation is reproduced in Appendix C.

A first important finding concerns the gap between general-market availability and institutional usability. As discussed in subchapter 2.3, Windows voice typing is a built-in speech-to-text function that can normally be used in many desktop applications, and in chapter 3.3 it was therefore included in the presentation as a basic point of reference for ordinary dictation tasks such as drafting an email or taking notes in Word. However, the preparation of the presentation showed that this general availability does not automatically translate into

institutional usability. In the context of preparing the presentation, the author contacted the DG TRAD Service Desk in her capacity as a proofreader in the European Parliament Romanian Unit. The email exchange with the DG TRAD Service Desk confirmed that Windows voice typing had been disabled on EP computers because it relies on a cloud feature and had been restricted by DG ITEC for security reasons (European Parliament n.d.b, Figure 9).

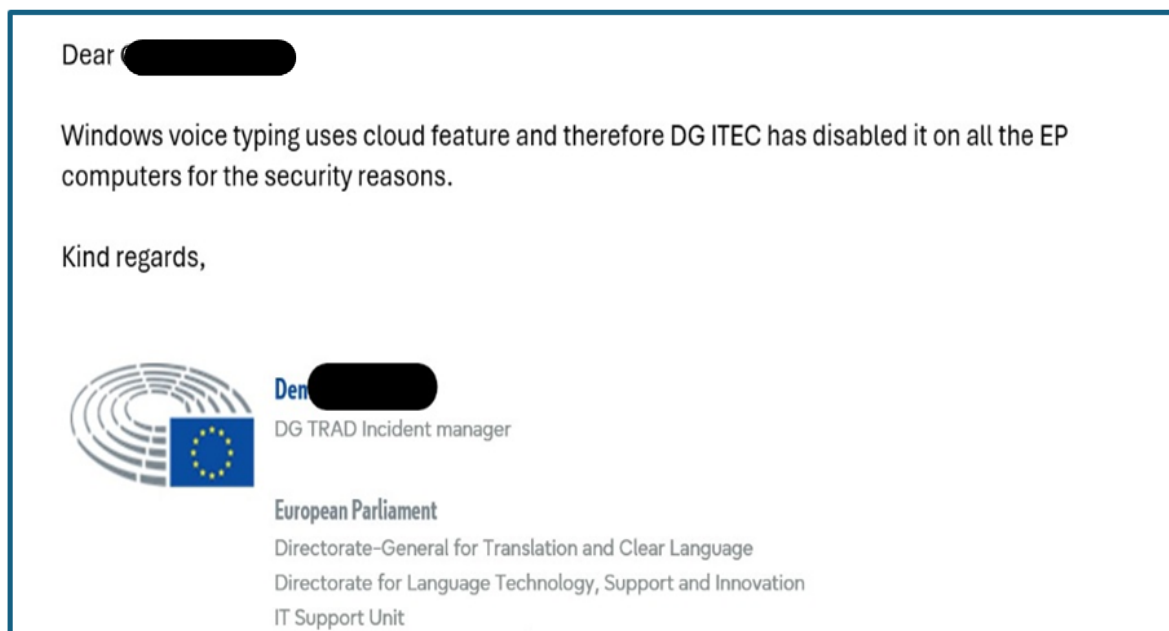


Figure 9: Email from DG TRAD Service Desk confirming Windows voice typing restrictions

([dgtrad.servicedesk@europarl.europa.eu](mailto:dgtrad.servicedesk@europarl.europa.eu))<sup>16</sup>

This finding is directly relevant to the distinction developed in chapter 2.3 between general-market ASR and institutional ASR. A tool may be technically mature and widely available, yet remain unusable in the institutional environment because governance and security requirements intervene before questions of convenience or efficiency even arise. It also sharpens the discussion in chapter 2.2 on workflow fit. In the theoretical foundation, ASR in translation was presented as potentially useful for small but frequent tasks such as drafting text, taking notes or supporting hybrid interaction with digital tools.

The preparation of the presentation showed that, in the European Parliament, this potential usefulness is filtered through institutional conditions. In practice, a widely used dictation function such as Windows voice typing cannot become part of everyday EP work if it is not available on institutional devices.

This means that relevance in theory and relevance in practice are not identical. Even where a tool could support a task efficiently, it cannot contribute to workflow integration if access to it is blocked at device level. The finding therefore supports the broader argument of this thesis that ASR adoption in institutional settings depends on governance, data protection and IT control.

<sup>16</sup> <https://www.europarl.europa.eu/trad/telework/>

A second important finding concerns Dragon Professional. As already discussed in chapter 2.4, Dragon Professional does not support Romanian, which already limits its relevance for Romanian-language professionals. During the preparation of the presentation, however, a further institutional layer became visible. The same email exchange with the DG TRAD Service Desk clarified that Dragon was not simply a tool that could be installed and used freely by any staff member. According to the information received, short-term licences of up to six months could be managed internally, whereas long-term or permanent licences were normally linked to medical justification reasons (European Parliament n.d.b, Figure 10).

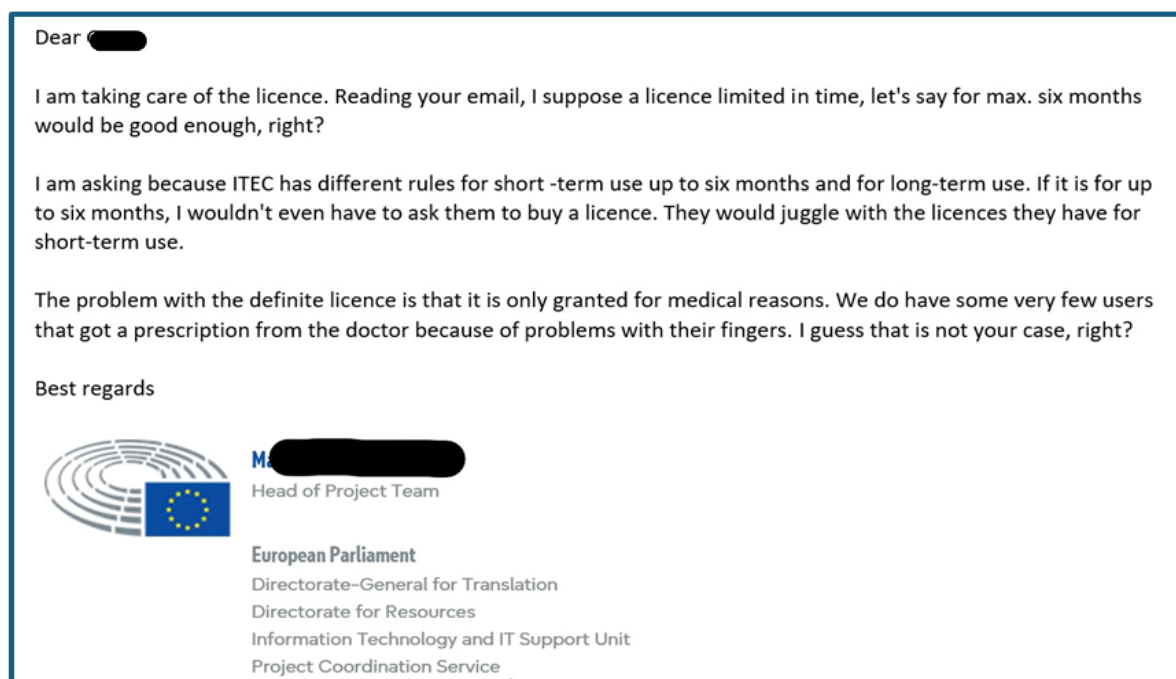


Figure 10: Email from DG TRAD Service Desk confirming different rules for short-term and long-term Dragon use

(dgtrad.servicedesk@europarl.europa.eu)

This finding adds a more concrete institutional dimension to the limitations already outlined in the theoretical foundation. In other words, Dragon is not only restricted linguistically for Romanian, but also organisationally through licensing conditions. It also confirms that, in the EP context, ASR is prioritised mainly in relation to accessibility and specific justified needs, rather than as a generally available tool for everyday translation work.

More importantly for the present study, the Service Desk response did not identify a generally available solution for active ASR use in Romanian. The only alternative suggested by the service desk was Dragon, yet for Romanian this did not amount to a practically available solution because the language itself is not supported (cf. Nuance 2023). At the same time, even for other languages supported by Dragon, access without a medical certificate remained limited to a temporary licence.

A further finding concerns the relative position of Trados Online Editor within this landscape. The preparation process showed that, among the tools examined for the presentation, Trados Online Editor remained the clearest public example of ASR directly integrated into a CAT-tool environment with Romanian-language support. This does not represent a finding about user attitudes yet, since these belong to the follow-up questionnaire, but it is still a

relevant finding from the preparation phase because it helped narrow down what could realistically be presented as a meaningful Romanian-language ASR in a CAT-tool example. In other words, while Windows voice typing was institutionally blocked and Dragon was limited both linguistically and institutionally, Trados Online Editor remained the most concrete example of what Romanian-language ASR in a CAT-tool setting can currently look like. This finding corresponds closely to the discussion in chapter 2.4, where Romanian support across tools was shown to be uneven (cf. Microsoft Windows n.d.; RWS n.d.; Nuance 2023), and to chapter 2.2, where the need to examine ASR in relation to revision, navigation and segment-based work was emphasised.

At the same time, this finding also points to an important limit. At the time of writing this thesis, Trados Online Editor is not used as part of the regular workflow of EP translators, whose work is based on the desktop Trados environment. It is therefore relevant here primarily as a publicly documented and illustrative example of Romanian-language ASR integrated into a CAT-tool interface, not as evidence of current institutional practice. Nevertheless, because the EP workflow is based on the Trados Studio desktop environment, the online version remains useful as an indication of functions that could become relevant in the future, should similar ASR and voice-command features also be introduced into the desktop workflow.

The preparation of the presentation therefore did not reveal any generally available solution for active ASR use in Romanian under ordinary EP working conditions. Instead, it revealed a partial and uneven landscape: Windows voice typing was institutionally blocked, Dragon was the only alternative proposed by the service desk but did not support Romanian and was only temporarily available without medical justification, and Trados Online Editor remained a public example rather than a tool embedded in the actual EP workflow. This result is fully in line with the broader thesis argument that ASR in institutional settings must be examined not only in terms of technical capability, but also in terms of actual workflow compatibility under organisational constraints.

Overall, the findings resulting from the preparation of the presentation make a specific contribution to RQ2: they show that, at the time of writing, active ASR use cannot yet be considered a practically available option for Romanian-language translation work in the European Parliament under ordinary institutional conditions. ASR is already present in the EP in several forms, but with different functions and degrees of accessibility: Verba as an embedded internal transcription tool, the Speech-to-Text platform as a public live ASR service for accessibility, and Dragon as a tool available only under specific conditions, including accessibility-related needs or limited testing periods. However, none of these currently provides a generally available Romanian-language solution for active dictation and voice-enabled interaction in the everyday desktop translation workflow. In this respect, the preparation of the presentation points less to current workflow integration for Romanian than to a gap between technically demonstrable functionality and institutionally available practice. At the same time, Trados Online Editor still remains relevant as an indication of what future development might look like if similar ASR and voice-command features were later introduced into the desktop CAT-tool environment on which the EP workflow is currently based.

This empirical step also helped address one of the gaps identified in the theoretical foundation, namely the lack of concrete information on workstation-level ASR use in the European Parliament. By documenting the restrictions discovered during the preparation of the



presentation, it offers a more grounded view of what institutional constraints mean in practice. At the same time, the presentation also had a second function: it contributed to spreading more knowledge about ASR within the Romanian Translation Unit of the European Parliament. In particular, it introduced Dragon as a less familiar ASR option that may be relevant for accessibility-related use or for testing over a limited period.

Further, this empirical step also has clear limitations. The findings are based mainly on the preparation of one presentation, on the checking of a limited number of ASR options and on information received through DG TRAD Service Desk communication. They do not amount to a full technical audit of ASR availability across the institution and do not include systematic testing across multiple departments, devices or user groups. Further, this subchapter does not yet address how Romanian-language professionals evaluated these tools after the presentation; this question is treated separately in the follow-up questionnaire findings. Further work could therefore have included broader technical verification across EP devices, more detailed comparison between institutional and non-institutional setups, or direct consultation with additional IT and workflow stakeholders. Even so, despite these limitations, the preparation of the presentation makes a useful contribution to the thesis because it helped prepare the Romanian Unit audience for the questionnaire by spreading more knowledge about ASR and by illustrating how ASR may function in translation-related workflows. At the same time, it clarified some of the institutional constraints surrounding workstation-level ASR use and suggested that ASR could in future become more present in CAT-tool-related work under EP conditions.

#### **4.4 Short findings from the follow-up try-out**

Although the follow-up try-out was not designed as a fully separate data-collection stage, it nevertheless provided a small number of observations that are useful for interpreting the questionnaire results. A first relevant point is that five respondents chose to remain after the presentation and test some of the demonstrated ASR solutions despite limited time and availability. This suggests that interest in ASR was not confined to a purely abstract or theoretical level. At least for some participants, the topic appeared relevant enough to justify immediate practical engagement with the tools.

The try-out itself remained brief and exploratory. Participants did not carry out a full translation or post-editing task in either the desktop or the online environment. Instead, they tested a limited number of representative actions that reflected the functions demonstrated in the presentation. These included dictating a segment from scratch, moving between segments through voice commands and revising pretranslated material by using commands for punctuation or deletion. In this sense, the session did not aim to generate a separate body of findings comparable to those derived from the questionnaire, but rather to provide a more concrete practical frame of reference for participants before they expressed their views in written form.

Even within these limits, the try-out was useful because it showed that direct contact with the tools encouraged participants to think more specifically about their possible relevance for professional practice. Rather than responding only to the idea of ASR, participants were able to reflect on actual interaction with selected functions, however briefly. This helped transform

initial curiosity into more concrete consideration of questions such as usefulness, correction effort and compatibility with institutional workflows.

At the same time, this part of the empirical material has clear limitations. Only five respondents took part in the follow-up try-out, and participation depended on individual availability after the presentation. This means that the session involved a small and self-selected group and cannot be treated as representative of the wider Romanian Unit, composed at the time of 22 translators, proofreaders and the Head of Unit. In addition, the try-out remained exploratory rather than systematic. Participants did not perform a complete translation or post-editing task in either Trados Studio desktop or Trados Online Editor, but only tested a small number of selected functions. The observations therefore reflect immediate practical reactions rather than sustained use under real working conditions. Another limitation is that the session was not designed as a separate structured data-collection stage, so no comparable performance data were gathered and no uniform testing protocol was applied across all participants. A more systematic design could have included a larger number of participants, a scheduled hands-on session with the same tasks for everyone and longer testing in both the desktop and online environments. Even so, the try-out remains relevant for the present thesis because it offered a first practical indication of participant interest and created a more concrete basis for the questionnaire responses.

#### **4.5 Findings from the questionnaire (EP RO Unit, DG TRAD)**

The questionnaire offers the most direct view of how the presentation and, for some participants, the optional follow-up try-out were received by Romanian-language professionals of the European Parliament.

As outlined in subchapters 3.4 and 3.5, this part of the methodology was designed to capture perceptions of ASR in relation to relevance, usability and acceptance after participants had first been introduced to a range of concrete examples and possible use cases of ASR and, where they wished, had the opportunity to engage with the tools more directly in the follow-up try-out. In this way, the questionnaire adds a more personal and more systematic layer to the preceding findings. It contributes primarily to RQ3, namely how Romanian-language professionals of the European Parliament perceive and experience ASR technologies. The questionnaire and the compiled results are included in Appendix D.

The questionnaire was completed by nine respondents, all of them translators or proofreaders from the EP Romanian Unit. As noted earlier, the unit consisted at the time of 22 staff members, including translators, proofreaders and the Head of Unit. The questionnaire therefore reflects the views of 9 out of 22 Unit members and should be read as a small-scale institutional snapshot rather than as a representative survey. Most respondents had little previous experience with ASR. Before the presentation, four said in the questionnaire that they had never used ASR technologies, three had used them only occasionally and two said that they used them sometimes, for example for notes, emails or in other private-life situations. This point is important for the interpretation of the whole questionnaire. The later positive answers do not mainly reflect the views of participants who were already very familiar with speech technologies. On the contrary, they show that even respondents with little or no previous

experience were able to recognize possible advantages of ASR after a structured explanation and a practical demonstration.

This becomes visible already in the first question, which asked whether the presentation helped participants understand better what ASR is and what it can be used for. Eight of the nine respondents selected the highest value on the scale and the remaining respondent selected the second-highest value. The result is therefore almost unanimous. It suggests that the presentation succeeded in creating a clearer understanding of the topic before participants were asked to evaluate it more critically. This is important methodologically, because the questionnaire was not designed to collect opinions in the abstract, but to collect opinions after participants had first been introduced to the concept through examples and demonstrations. As described in chapter 3.3, the presentation moved from familiar situations to more specific institutional and translation-related uses of ASR. In this sense, the questionnaire also shows the effect of the presentation itself in making the topic more understandable and relevant.

The next two questions show that this improved understanding was followed by a clearly positive view of ASR's possible usefulness. When participants were asked whether the examples shown in the presentation seemed relevant to their professional activity, eight selected the highest value and one selected the middle value. No respondent chose a negative answer. This suggests that the examples were not seen as abstract or too disconnected from the day-to-day work of the Romanian Unit. Rather, participants were generally able to see ASR as relevant for concrete professional tasks. The responses to the next question are even stronger: when asked whether ASR seemed useful in private life as well, eight respondents again selected the highest value and one selected the second-highest value. Taken together, these two questions show that participants located the value of ASR in both spheres: professional and private. This is in line with the presentation strategy: ASR was first shown in familiar everyday contexts and then in work-related environments such as Word, email and CAT tools. This progression seems to have helped participants see ASR not only as an existing technology, but as one that could also fit certain professional tasks.

The evaluation of the specific tools shown during the presentation follows the same positive pattern. The Windows dictation function was considered easy to use by all nine respondents: seven selected the highest value and two the second highest. Dragon NaturallySpeaking received similarly positive evaluations, with six highest ratings and three second-highest ratings. This last favourable evaluation should, however, be interpreted in relation to the way the tool was presented and tested: as an English-language professional desktop benchmark rather than as a currently supported Romanian-language solution for the EP desktop workflow. The same applies to the dictation and voice-command functions shown in Trados Online Editor, which also received six highest and three second-highest ratings. These responses suggest that the demonstrated tools were not perceived as too difficult, too technical or too unrealistic for work-related tasks in general, including translation-related tasks. From the users' perspective, ASR solutions appeared understandable and potentially usable in practice. This finding is especially meaningful when read together with chapter 4.3, where the institutional limits of such tools were discussed. As shown there, the main problem was not that the tools themselves seemed too complex or unattractive, but that their availability under EP conditions is uneven for different languages or restricted in general. The questionnaire therefore adds an important user-centred perspective: the participants did not reject these solutions. On the contrary, their

responses suggest that the tools were generally perceived positively, while the main limits lie in the institutional context rather than in user attitudes.

A more nuanced picture appears in the question about pronunciation, pauses in speech and accent. This was the most mixed question in the questionnaire. Three respondents selected the lowest value, one the second-lowest, three selected the middle value and two the second highest. No one selected the highest value. This spread of answers suggests that respondents did not perceive the issue in the same way. For some, pronunciation or accent did not seem likely to cause major difficulties. For others, these factors remained a clear concern. This finding corresponds closely to the theoretical discussion in subchapter 2.1, where speaker variability, pauses and pronunciation were described as factors that may affect recognition quality and correction effort. In addition, this result is relevant in relation to the methodology described in subchapter 3.5, since the participants included professionals with different accent profiles, including one non-native Romanian speaker, who also participated in the try-out. The responses therefore suggest that, while ASR was generally viewed positively, some uncertainty remains about how well it would work for different speakers in practice.

At the same time, this uncertainty did not lead to rejection of the technology. This becomes clear in the question that asked whether respondents would be willing to try an institutionally approved ASR solution in their day-to-day translation work. Eight respondents selected the highest value and one selected the second highest. This is one of the strongest results in the questionnaire. It shows that even those participants who had some doubts about pronunciation or errors were still largely open to trying such a tool. In terms of the analytical framework of the thesis, this is a strong indicator of acceptance in principle. It suggests that Romanian-language professionals in this group did not regard ASR as something inappropriate for their work as such. Rather, they were prepared to consider it, provided that the solution was suitable and approved for institutional use.

The multiple-choice questions make this openness more concrete by showing which types of ASR use seemed most relevant to the respondents. When asked which demonstrations appeared most useful, the options most often selected were Windows dictation and the possibility of integration into Trados Studio desktop, each chosen by eight respondents. Dragon NaturallySpeaking and Trados Online Editor dictation with voice commands were each selected by seven respondents, while the accessibility role of ASR was selected three times. This distribution suggests that respondents were primarily interested in workstation-level uses that could support their own daily tasks. Their strongest interest was therefore not in ASR as a broader institutional infrastructure for accessibility, but in tools that might help them write, revise, navigate or interact more efficiently with text and translation environments. This is an important contribution to the thesis because one of the gaps identified in chapter 2.5 concerned the lack of user-centred evidence on workstation-level ASR in the European Parliament. The questionnaire begins to fill that gap by showing practical interest among Romanian-language professionals

The same practical orientation appears even more clearly in the next multiple-choice question, which asked in which professional activities respondents would realistically find ASR useful. The most frequently selected option was notes or summaries, chosen by eight respondents, followed by drafting emails with seven selections. Post-editing pretranslated

segments with machine translation or TMX was selected by six respondents, while dictating a translation draft from scratch was selected by five. This pattern suggests that respondents did not associate ASR primarily with full dictation. Rather, they saw greater value in more limited and realistic workflow uses, especially for short support tasks and for work based on already available text. This result is important in relation to chapter 2.2, where the thesis argued that ASR in translation should be examined not only as a drafting tool, but also as a possible support for revision and post-editing. The questionnaire findings support that argument. They suggest that Romanian-language professionals may see ASR as fitting more naturally into hybrid workflows, in which speaking, correcting and navigating are combined, than into a fully speech-based translation process.

This pattern is also significant in relation to the limited hands-on exposure described in subchapter 4.4. While the try-out briefly familiarised participants with the main active uses of ASR, including segment-level dictation from scratch, navigation and revision, the questionnaire responses still point more clearly toward post-editing and support tasks than toward full draft dictation. This suggests that respondents related the demonstrated functions to their own work routines and saw the strongest potential in selective, task-specific uses rather than in replacing the translation act as a whole.

The question concerning the most likely context of use adds a further practical dimension. Five respondents answered that they would be most likely to use ASR both on their personal laptop and on the institution laptop. Three said mainly on their personal laptop or in private contexts such as at home or in the car, while only one selected mainly the institution laptop. This result suggests that respondents did not imagine ASR primarily as a tool confined to one context only. At the same time, the relatively low number of respondents who chose mainly the institution laptop also reflects the current institutional reality discussed in earlier chapters. Even when respondents were positive about the usefulness of ASR, they did not automatically assume that the institutional laptop was the place where it would most easily be used. In this sense, the answers reflect an awareness of the practical restrictions that currently shape technology use within the EP.

This becomes even clearer in the following question, which asked what seemed to be the biggest obstacle to adopting ASR in the institutional context. Seven respondents selected IT restrictions, the lack of an institutional cloud or the lack of institutional ASR technology, while only two selected confidentiality or security. This finding confirms from the users' perspective what had already emerged in subchapter 4.3. The main barrier is not a lack of interest in ASR, nor a general rejection of speech technologies, but the absence of suitable institutional infrastructure and approved technical solutions. Confidentiality and security remain important, but in this small sample they appear to be perceived more as part of the broader institutional framework than as the most immediate reason for non-use. In other words, the respondents seem to suggest that ASR could become more acceptable if the institution itself provided the necessary technical conditions in a secure and approved form.

The same cautious openness appears in the question asking whether respondents would be interested in trying ASR solutions, including the solutions illustrated in the presentation. Four answered yes and five answered maybe, while no respondent chose a negative answer. Although this is not an overwhelming yes, it is still a clearly positive result. It shows that

respondents were not only open to the idea of ASR in theory but were at least willing to consider a more practical step. As noted in subchapter 4.4, a small follow-up try-out did in fact take place after the presentation and five participants showed up. This further supports the interpretation that interest in ASR within the group was not merely abstract but also translated into a limited willingness to test the tools in practice when the opportunity was available.

The open-ended questions, although not mandatory, add an important qualitative layer to the questionnaire findings because they show more clearly how respondents themselves explained their answers. At the same time, the number of responses was lower than in the closed questions, which is to be expected in this type of questionnaire. Six respondents answered the question on the most useful idea or demonstration from the presentation, seven answered the question on the main advantages and disadvantages of ASR, seven answered the question on possible integration into the EP translation process and only two respondents added further optional comments. These answers are therefore limited in number, but they are still useful because they make the overall tendency of the questionnaire more concrete and show which aspects of ASR remained most salient for participants after the presentation and try-out.

In the first open question, which asked what seemed the most useful ASR idea or demonstration from the presentation and whether respondents already used speech recognition in private or professional life, the answers were quite varied. Some respondents referred to dictation functions in Windows applications or Trados Studio, while others pointed to speech technologies already familiar from everyday life, such as Siri or instant messaging on WhatsApp. One respondent explicitly stated that the most useful idea was the possibility of dictation in Studio or Windows. Another considered Siri voice assistance especially useful because it could be applied immediately in private life and without work-environment restrictions. One respondent highlighted dictating emails as the most useful demonstration, while also noting that the other examples were useful too. Another stressed the value of ASR in situations where one cannot use one's hands to type, for example while driving. Finally, one respondent referred explicitly to dictating and editing translations and stated that such a solution would be useful in professional life.

Taken together, these answers show that the presentation succeeded in linking ASR to recognizable situations, both private and professional. These open answers help refine the interpretation of the more positive responses in the previous questions. The fact that several respondents referred either to immediately usable private-life functions or to possible translation-related dictation suggests that ASR was interpreted as something that already fit into their existing private-life habits or could fit into specific institutional work tasks.

The next open question, on the main advantages and disadvantages of ASR, received seven answers and provides an even clearer picture of how respondents balanced the potential of ASR technology against its limits. On the positive side, several respondents mentioned saving time and improving efficiency. Another respondent pointed out that dictation can be easier than typing because it allows the user to speak instead of remaining constantly at the keyboard and, in some situations, to carry out another activity at the same time. This was also understood as a practical gain in efficiency. One answer referred more broadly to productivity, accessibility and the possibility of combining ASR with other AI technologies, as well as to the automatic processing of larger volumes of spoken content, such as meetings or conferences.

All these answers suggest that respondents mainly associated the advantages of ASR with practical support in everyday work: saving time, improving efficiency, reducing the need for constant keyboard input. This corresponds closely to the discussion in chapter 2.2, where ASR was linked both to possible efficiency gains and to ergonomic benefits, such as a more comfortable working setup and less continuous dependence on keyboard and mouse input (cf. Ciobanu & Secară 2019: 97–101).

One answer is especially interesting in relation to translation-specific workflows. A respondent noted that saying a translation aloud may help in finding better translation solutions. This point goes beyond the idea of ASR as a simple dictation tool. It suggests that speech input may also influence the translation process itself, for example by supporting drafting, reformulation or self-revision. In this sense, the open answers suggest that some respondents saw ASR not only as a technical aid for entering text, but also as a possible support for the cognitive side of translation work.

At the same time, the disadvantages mentioned in these open answers are equally revealing. Several respondents referred to recognition errors, especially the possibility that the software may replace one word with another or transcribe badly and require much more correction afterwards. One answer referred specifically to difficulties caused by accent or pronunciation. Others mentioned confidentiality, data protection and security. One respondent also pointed to linguistic or cultural bias. These concerns correspond very closely to the issues discussed in subchapters 2.2, 2.3 and 2.4. On the one hand, they reflect the usability dimension of the thesis, since recognition errors and post-correction effort directly affect whether ASR is worth using in practice. On the other hand, they reflect the acceptance dimension, since confidentiality, security and institutional compliance are especially important in the EP context. The open answers therefore reinforce the interpretation already suggested by the closed questions: the respondents' attitude was generally positive, but not uncritical. They clearly identified potential advantages, yet they also remained aware of the limits that could reduce ASR's usefulness or make its broader institutional use problematic.

The following open question asked whether respondents thought ASR could be integrated into the translation process within the European Parliament and why or why not. This question also received seven answers, and the overall tendency was again positive or conditionally positive. Some respondents simply answered yes, without further explanation, which in itself is meaningful because it suggests that they did not regard integration as implausible. Others gave more specific reasons. One respondent stated that ASR could be helpful in certain medical situations, as well as in the stages of translation, proofreading and self-revision. Another wrote that ASR could be integrated for most EP documents because the technical solutions already exist and proper encryption and secure cloud storage are possible. One respondent considered that integration would be possible but would require some investment in cloud infrastructure. Only one respondent expressed a more clearly sceptical position, stating that they did not think such integration would happen very soon because of confidentiality concerns.

These answers suggest a more precise interpretation than a generally positive attitude toward ASR. Read together with the earlier questionnaire results, they show that the respondents did not view integration into the EP translation process as impossible in principle, but neither did they imagine it as something that could simply be introduced under the current institutional

conditions. Rather, their answers indicate that possible integration is seen as dependent on an approved technical and organizational framework: secure infrastructure, suitable cloud or storage solutions, encryption, and clear confidentiality safeguards. This interpretation also corresponds to the broader pattern already identified in this chapter. The respondents evaluated the demonstrated tools positively, but at the same time identified IT restrictions, the lack of institutional cloud solutions and the lack of approved institutional ASR technology as the main obstacles. Their answers therefore point not to a rejection of ASR as such, but to a distinction between technical potential and institutional feasibility. In this sense, Romanian-language professionals in this group appear broadly open to selective and workflow-specific forms of ASR use, especially in hybrid tasks such as translation, proofreading, self-revision, note-taking or post-editing, but only insofar as these can be integrated into the EP environment in a secure, approved and practically workable way.

The final optional comments were answered by only two respondents, so they cannot be treated as representative. Even so, they are worth mentioning because both were clearly positive and because they comment not only on ASR in general, but also on the presentation itself. One respondent described ASR as a very useful topic and stated that, although aware of its existence before, they had not realized how useful it could be both privately and professionally. The same respondent also considered the presentation very well structured and the examples relevant. Another respondent congratulated the presenter and expressed the hope that ASR would soon also be integrated into professional life. These comments support the interpretation that the presentation itself helped create a more informed and more favourable view of ASR.

All these questionnaire findings make it possible to answer RQ3 more clearly. They suggest that Romanian-language professionals in this small EP group generally perceive ASR as relevant, potentially useful and worth trying, especially after being introduced to concrete examples and, in some cases, after a short hands-on try-out. The strongest positive reactions concern workstation-level and hybrid uses, such as note-taking, drafting emails, post-editing and possible CAT-tool interaction, rather than a fully speech-based workflow. In this sense, the questionnaire complements the previous findings by showing that possible users do not reject active ASR use as such. On the contrary, they are broadly open to it and able to imagine concrete professional uses for it. At the same time, the findings also show that this openness remains conditional. The respondents' acceptance is shaped by concerns about pronunciation, recognition errors, correction effort and, above all, institutional constraints such as IT restrictions, lack of approved infrastructure, confidentiality and data protection. In this way, the questionnaire helps cover one of the gaps identified in the theoretical foundation, namely the lack of Romanian-language, workstation-level user feedback on ASR under institutional conditions. It also supports the broader argument of this thesis that the relevance, usability and acceptance of ASR depend not only on the existence of the technology itself, but on how well it fits actual workflows and institutional requirements. Read together with subchapters 4.1-4.3, the questionnaire also shows that the main barrier to wider uptake is currently institutional implementation rather than user resistance.

Overall, the questionnaire results place the Romanian respondents between openness to innovation and a strong awareness of institutional constraints. The answers do not point to a contradiction between relevance and caution. Rather, they show that both can coexist: participants may recognise the potential usefulness of ASR for note-taking, email drafting and



even self-revision, while at the same time judging actual adoption to depend on secure infrastructure, approved tools and realistic Romanian-language support. This distinction helps explain why positive evaluations at user level do not automatically lead to routine uptake at workstation level. In the EP Romanian context, the missing link is therefore not basic user interest, but the alignment between technical availability, institutional approval and everyday workflow fit. Read in this way, the questionnaire does not simply show general enthusiasm for ASR. It shows a more mature form of acceptance: participants are open to the technology, but mainly under conditions that would make its use professionally manageable and institutionally legitimate.

At the same time, this part of the empirical material has clear limitations. The number of respondents is limited, so the results cannot be treated as representative of all Romanian-language professionals in the European Parliament or of other European institutions. The questionnaire captures reactions after one presentation and an optional short try-out, rather than long-term experience with regular use. In addition, the open-ended questions were not mandatory, which means that the qualitative comments remain limited in number. The findings are therefore based mainly on self-reported perceptions and expectations, not on direct measurements of correction effort, productivity gains or long-term adoption. Further research could have included a larger sample, more systematic practical testing and closer observation of how Romanian-language professionals actually use ASR over a longer period.

Even so, the questionnaire makes an important contribution to the present thesis because it adds the perspective of potential users themselves and shows that, in this Romanian-language institutional setting, the main obstacles to wider uptake are not primarily attitudinal, but practical and institutional.

## **5. Concluding observations**

This final chapter brings together the main arguments and findings of the thesis and returns to the central question that has guided the study from the beginning: how relevant, usable and acceptable are current automatic speech recognition (ASR) technologies in European institutional settings, with particular attention to Romanian and to the European Parliament?

The preceding chapters have shown that this question cannot be answered in a simple affirmative or negative way. The value of ASR depends strongly on the form in which it is used, the workflow into which it is introduced, the linguistic conditions under which it operates and the institutional environment that either enables or restricts its deployment.

The theoretical foundation showed that ASR has evolved from an experimental technology into a broad range of tools used for dictation, transcription, speech-enabled interaction and multilingual institutional services. At the same time, the theoretical foundation of this thesis made clear that technological progress does not automatically lead to professional uptake. In institutional and translation-related settings, the key issue is not only whether ASR works technically, but whether it fits concrete tasks, keeps correction effort manageable and can be

used within requirements related to confidentiality, data protection and IT governance. These considerations form the basis of the present study and guided the analysis of ASR in a concrete institutional setting, with particularly attention to the Romanian language.

Against this background, the empirical part of the thesis examined three forms of ASR use in the European Parliament: live transcription, asynchronous transcription and active dictation together with voice-enabled interaction. The mixed-methods design made it possible to compare an expert perspective on the European Parliament's live Speech-to-Text platform, a user perspective on Verba, DG TRAD's internal ASR tool, and the perspectives of Romanian-language professionals on workstation-level ASR following a presentation, a short try-out and a questionnaire. Rather than seeing ASR as a technology with one general use, the thesis approached it as a language tool that takes different forms in different workflows, with relevance, usability and acceptance varying according to function, risk and institutional context.

The purpose of the following subchapters is to interpret the results more broadly and to show more clearly how they contribute to the existing literature and to the Romanian European institutional context. Subchapter 5.1 revisits the main findings in relation to the gaps identified in the theoretical foundation and shows how the empirical material answers the research questions. Subchapter 5.2 reflects on the strengths and limitations of the study and clarifies in what sense this thesis makes a useful contribution. Subchapter 5.3 then looks forward and considers the practical implications of the findings, together with recommendations for future implementation and further research. In this way, the conclusion does not simply summarize the thesis but draws out its broader significance for understanding ASR in multilingual institutional environments, with regard to the Romanian language and community.

## **5.1 Summary of key findings in relation to the literature**

The findings of this thesis should be read in direct relation to the problems and gaps identified in the pre-existing research and institutional documentation reviewed in chapter 2 and to the methodological choices explained in chapter 3.

More specifically, chapter 2 showed that ASR has developed into a mature and varied technology, but also that technical progress alone does not explain its relevance, usability or acceptance in institutional settings. This is precisely the perspective adopted in the present thesis, which examines these three dimensions through a case study on the Romanian language in European institutions. On this basis, chapter 2 identified three main gaps: limited transparency regarding how institutional ASR is developed and evaluated in practice, insufficient attention to the fact that ASR does not have the same value across different workflows and a lack of Romanian-focused evidence from institutional settings. Chapter 3 responded to these gaps through three research questions and a mixed-methods design based on interviews, a presentation, a try-out and a questionnaire.

A first major gap concerned the limited transparency of institutional ASR. Public institutional documents confirmed the existence of tools such as the European Parliament's live Speech-to-Text platform and Verba (cf. European Parliament, Directorate-General for Translation 2024: 34), but they did not explain clearly enough how these tools function in

practice, how they are adapted to parliamentary needs or how they are experienced by professional users. The two interviews, one with the Head of the Speech-to-Text Unit, Mr. Dragoș-Lucian Ciobanu and one with an anonymous experienced Romanian Verba user from the Directorate-General for Translation, helped reduce this gap from complementary angles.

The interview with the Head of the Speech-to-Text Unit first confirmed that the live platform is conceived primarily as an accessibility-oriented service whose institutional justification lies in widening access to parliamentary communication. This corresponds closely to the theoretical discussion of ASR as multilingual infrastructure for accessibility and participation, but it also makes that discussion more concrete by showing that, in the European Parliament, live ASR is not presented as a substitute for interpreting or for human linguistic mediation in general.

At the same time, the interview showed that accessibility is not the only source of relevance for the live platform. It is also useful as a support tool in a multilingual institution where several events may take place simultaneously and where Members, assistants and staff may need rapid written access to ongoing discussions across languages. Institutional reporting often presents live ASR mainly through the lens of public accessibility (cf. Mavrič 2022: 35–36; European Parliament, Directorate-General for Translation 2024: 34), but the interview suggests that it also serves internal needs related to monitoring, multitasking and the rapid recovery of information in multilingual event environments. In this respect, the findings refine the literature by showing that institutional relevance is broader than public accessibility alone.

A further contribution concerns the institutional architecture of the Parliament’s live ASR solution. The same interview made clear that the current platform is not a purely internal, closed, end-to-end system. Rather, it combines open and institutional infrastructure with separate modules for language identification, transcription and translation, while also relying on a commercial translation component under controlled institutional conditions. The theoretical foundation of this thesis emphasised initially the tension between institutional control and commercial cloud-based technologies. The findings do not remove that tension, but they show that the decisive distinction is not simply between “institutional” and “commercial”, nor between “local” and “cloud-based”, but between uncontrolled services and components that can be integrated into an ASR governed institutional environment with clear rules on access, data handling and customisation. The interview also confirmed that domain adaptation remains essential. General-language ASR is not sufficient for parliamentary speech, where terminology, acronyms, names and institutional phraseology strongly influence output quality. In this sense, the thesis supports the literature on domain-specific adaptation (cf. Di Nuovo 2023: 185–195), while also giving it more practical institutional insight. The relevance and usability of institutional ASR do not depend only on general recognition quality, but also on how well the system is further customized by professionals to the setting in which it is expected to function.

Another important finding from this interview is that the development of institutional ASR is constrained not only by technical questions, but also by very practical institutional limits. The interview with the Head of Unit pointed to human-resource pressures, budgetary restrictions, uneven language support and difficulties in conducting systematic multilingual evaluation. This helps explain why transparency and detailed benchmarking are often missing from public institutional reporting. The absence of detailed evaluation is not simply a matter of insufficient public communication. It is also linked to the fact that multilingual quality checking

across 24 languages require a large amount of coordinated human work. This gives a more realistic explanation of the transparency gap identified in the literature.

The same applies to data protection and access control. Earlier chapters stressed that institutional ASR must function within confidentiality and governance requirements. The interview confirmed this directly by showing that live transcripts cannot always be stored and re-checked afterwards, especially in more politically sensitive or non-public contexts. This means that data-protection rules affect not only deployment, but also the evaluation and further improvement of the live ASR platform.

A second major literature gap concerned the need to examine ASR as a workflow-dependent technology whose value changes according to the function it serves. One of the clearest conclusions of the thesis is that live ASR, asynchronous transcription and active workstation-level ASR involve different expectations, different risks and different acceptance thresholds. Live ASR is valuable because it delivers access immediately, but that same immediacy also makes it vulnerable in politically sensitive settings, where a single error may have reputational consequences, as confirmed in the interview with the Head of the Speech-to-Text Unit. Asynchronous ASR, by contrast, is more acceptable precisely because human revision remains possible before publication on official EU websites.

The contrast between the live platform and the internally developed Verba of DG TRAD is particularly revealing in this respect. The interview with the Head of the Speech-to-Text Unit showed that the live platform is not preferred for plenary use when the political risk of uncorrected instantaneous output is too high. The interview on Verba, however, showed that a more controlled transcription workflow can be highly relevant and broadly acceptable for the transcription of high-level meetings such as plenary sessions, precisely because correction and publication are separated in time. In other words, the same institution can accept ASR strongly in one setting and more cautiously in another, not because it changes its general opinion about the technology, but because the consequences of errors differ from one use case to the next.

The Verba interview also helps fill an important user-centred gap in the literature. Public institutional reporting confirms that Verba replaced CRE and functions across all official EU languages (European Parliament, Directorate-General for Translation 2024: 34), but it provides only limited insight into how the tool is experienced by those who actually work with it within DG TRAD. The interview with the anonymous Romanian proofreader shows that, from a user perspective, Verba is no longer a marginal convenience, but an established part of the workflow. Returning to manual transcription is experienced as highly demanding. This is a strong indicator of relevance in practice. The findings therefore support the idea already present in the literature that ASR becomes professionally meaningful when it reduces effort rather than adding to it. In the case of Verba, its value lies not in removing the need for human work, but in reducing significantly the amount of work required before publication.

At the level of usability, the same interview adds several useful refinements to the EP reporting gap. First, it suggests that Romanian output in Verba is generally manageable and confirms that the newer system is clearly experienced as an improvement over the previous CRE tool. Second, it shows that the amount of correction still depends strongly on speaker behaviour, pronunciation and speech organisation, which corresponds closely to the theoretical discussion of disfluency, variability and realistic speech conditions. Third, it suggests that

Romanian-specific orthographic issues, highlighted in subchapter 2.4 (Avram *et al.* 2020: 397–398), do not appear to be the most visible practical problem in the Verba workflow examined here. From the user perspective, the newer system is perceived as already delivering partly processed and normalized output, which may explain why earlier orthographic concerns seem less salient in everyday use.

This can be related more broadly to the theoretical discussion in subchapter 2.1 on the evolution of ASR towards more adaptable systems, while remaining cautious about Verba's internal architecture, which is not transparent from the user perspective. What emerges more clearly in practice are difficulties linked to newer parliamentary terminology, proper names, institutional designations and speaker-dependent variation. The findings therefore refine the Romanian-focused literature by suggesting that, in the Verba workflow examined through this interview, earlier orthographic concerns no longer appear to be the most visible practical issue from the user perspective, while terminology management and speech variability emerge more clearly as the main remaining difficulties.

At the time of writing, the literature reviewed offered only limited Romanian-specific evidence from within a European institutional ASR workflow. The interviews conducted for this thesis suggest a more differentiated picture. According to the Speech-to-Text Unit, Romanian appears to perform relatively well in the live EP setting, while the interview with the Romanian proofreader indicates that Romanian output in Verba is generally manageable in practice. At the same time, both interviews show that this positive picture remains incomplete. Romanian has not yet been examined through a detailed and systematic institutional error analysis, and the empirical material gathered in this study remains stronger in qualitative than in quantitative terms. The thesis therefore helps reduce the Romanian evidence gap by providing qualitative evidence that Romanian ASR can function meaningfully in these institutional contexts, while also showing that the gap has not yet been fully closed through specific quality evaluation.

A third major contribution of the thesis concerns the literature on ASR in translation and CAT-tool environments. Chapter 2 showed that the professional translation literature often focuses too narrowly on dictation from scratch, whereas real translation practice is usually hybrid and often centred on revision, post-editing and selective task support. The findings from the presentation, try-out and questionnaire strongly support this point. The most revealing result is that Romanian-language professionals did not see ASR primarily as a tool for replacing the entire translation act with speech. Their strongest interest was directed towards note-taking, drafting short texts such as emails, revising pretranslated segments and other selective workstation-level functions. Although dictation from scratch was presented as a possible use during the try-out session, it did not emerge as the dominant preference, even though some participants also expressed interest in it.

This finding is particularly relevant to the literature on the active use of ASR in translation environments. Earlier studies suggest that ASR is rarely used as a fully speech-driven mode of translation and is more often combined with keyboard and mouse in a hybrid workflow (cf. Ciobanu & Secară 2019: 97). The present study confirms and refines this insight in the Romanian institutional context. More specifically, the participants did not show the strongest interest in dictating entire translations from scratch. Instead, they appeared more interested in

selective and practical uses of ASR, such as revision, drafting short texts and other support tasks that can be integrated into existing ways of working. This suggests that, in institutional translation settings, ASR is currently more attractive as a complementary tool than as a replacement for the translation process itself. In this respect, the thesis contributes to a more realistic and practice-oriented understanding of how ASR may be adopted in professional translation workflows.

At the same time, the findings from the presentation and follow-up process sharpen another important point already present in the theoretical foundation: technical availability on the market does not equal institutional availability in practice. The investigation of workstation-level options showed that Windows voice typing, although broadly available in principle, is restricted on EP-managed laptops, Dragon Professional is limited both linguistically in relation to Romanian and organisationally through licensing conditions, and Trados Online Editor remains an illustrative public example rather than part of the actual EP translation workflow. This directly addresses the second research question and also fills a gap left open by the literature. Discussions of CAT-tool ASR integration often ask what is technically possible, but not always what is institutionally accessible. The findings of the present thesis show that this distinction is decisive. A tool may be functional in theory and still remain irrelevant for everyday European institutional work if it cannot be used within the approved technical environment.

Taken together, the presentation and follow-up process suggest that the central issue is not whether ASR tools exist or whether users are open to them, but whether they can be integrated into the institutional environment in a secure, approved and practically workable way. In this respect, the findings show that, at the time of writing, active ASR use cannot yet be regarded as a practically available option for Romanian-language translation work in the European Parliament under ordinary institutional conditions. ASR is already present in the EP in several forms, but these differ in function and degree of accessibility: Verba exists as an embedded internal transcription tool, the Speech-to-Text platform as a live public-facing accessibility service, and Dragon as a tool available only under specific conditions, such as accessibility-related use or limited testing periods. However, none of these currently provides a generally available Romanian-language solution for active dictation and voice-enabled interaction in the everyday Trados Studio desktop translation workflow.

In this respect, the preparation of the presentation points less to actual workflow integration for Romanian than to a gap between technically demonstrable functionality and institutionally available practice. At the same time, Trados Online Editor remains relevant as an indication of what future development might look like if similar ASR and voice-command functions were later introduced into the desktop CAT-tool environment on which the current EP workflow is based.

This pattern regarding the problematic institutional integration of ASR is also one of the clearest results of the last methodological step, the questionnaire distributed within the Romanian Translation Unit of EP DG TRAD. The respondents did not reject ASR because it seemed unattractive or too difficult to use. On the contrary, the demonstrated tools were generally evaluated positively in terms of apparent ease of use and potential usefulness. The main obstacles identified were not attitudinal, but structural: IT restrictions, limited institutional

infrastructure and the absence of approved Romanian-language solutions for everyday workplace use. This is one of the most important findings of the thesis, because it suggests that the main barrier to wider uptake is not user resistance, but institutional implementation.

At the same time, the questionnaire adds an important nuance regarding acceptance. Although respondents expressed mixed views about the possible influence of pronunciation, pauses and accent on ASR performance, they still showed a clear willingness to try an institutionally approved solution. This points to a form of conditional acceptance. Users do not expect perfection before engaging with the technology, but they do expect it to be secure, available, officially approved and sufficiently adapted to their professional environment. In this sense, acceptance is not simply a matter of personal enthusiasm, but of workflow fit and institutional legitimacy.

The qualitative comments support further this interpretation of the questionnaire results. Respondents linked ASR to time saving, easier drafting, productivity and possible ergonomic benefits, but also raised concerns about recognition errors, correction effort, confidentiality and security. In this way, their remarks confirm the same overall pattern: openness to ASR exists, but it remains closely tied to institutional and practical conditions. The comments show that participants do not reject the technology itself; rather, they see its professional value as depending on secure access, approved solutions, workable Romanian-language support and a realistic fit with everyday tasks.

This helps explain why positive attitudes towards ASR do not automatically translate into routine workstation use. What emerges is therefore not simple enthusiasm, but a cautious and conditional form of acceptance. In other words, the questionnaire identifies perceived advantages and obstacles, but also because it shows that workstation-level ASR in the Romanian EP context is already imaginable and professionally intelligible to users, even if it is not yet routinely available in practice.

Taken together, the findings answer the three research questions in differentiated ways.

- *RQ1: How are current institutional ASR solutions in the European Parliament described and evaluated by experienced professionals?*

Current institutional ASR solutions in the European Parliament are described by experienced professionals in relation to their specific functions, whose value depends on the role they play in a given workflow. The live Speech-to-Text platform is understood primarily as an accessibility service, but also as a support tool for multilingual monitoring and rapid access to spoken content in a complex institutional environment.

Verba, by contrast, is described as an established internal transcription tool that reduces effort in a workflow where human revision remains possible before publication. Their evaluation is therefore broadly positive, but it is also shaped by important limiting factors, including domain adaptation, confidentiality requirements, various resource constraints and the practical difficulty of carrying out systematic multilingual quality assessment.

- *RQ2: How can ASR be integrated into workflows in the European Parliament, and more specifically into the CAT-tool workflow, under organisational constraints?*

At the time of writing, Dragon Professional appears to be the only identifiable option for active ASR use in the EP environment, and even this is available only under specific conditions, such as accessibility-related use or limited licences.

Active ASR is therefore not yet part of the ordinary Romanian-language desktop translation workflow of EP DG TRAD. The findings nevertheless suggest that such integration is technically conceivable and acceptable, but only if an institutionally approved Romanian-language solution is made securely available, supported within the EP technical environment and adapted to the actual needs of the workflow.

- *RQ3: How do Romanian-language professionals in the European Parliament perceive and experience ASR technologies?*

Romanian-language professionals in the European Parliament perceive ASR in differentiated ways depending on whether they evaluate an existing institutional tool or the possibility of broader future use. The experienced Verba user describes Romanian ASR as manageable and useful within an existing transcription workflow, although its performance still depends on factors such as speaker behaviour, terminology, names and speech organisation.

The questionnaire respondents, by contrast, perceive ASR more generally as potentially useful and increasingly relevant, especially for selective support tasks such as drafting, note-taking, revision and post-editing rather than for replacing the translation process as a whole. Their overall attitude is therefore broadly open, but this openness remains conditional on secure access, institutional approval and a realistic fit with everyday professional practice.

The broader conclusion emerging from these findings is that the relevance, usability and acceptance of ASR in European institutions must be assessed in relation to function, workflow and institutional conditions, rather than against a single general standard. An ASR tool is relevant when it responds to a clear professional need within a specific workflow. It is usable when its output is sufficiently reliable for the task in question and the remaining correction effort is manageable in practice. It is acceptable when it is not only technically functional, but also secure, institutionally approved and compatible with the professional environment in which it is intended to be used.

In this sense, the case of Romanian in the European Parliament shows that ASR is no longer absent from the institutional landscape, but neither is it yet fully integrated across all relevant workflows. Romanian is already present in important forms of institutional ASR use, especially in live accessibility services and internal transcription workflows, where it appears to function in a generally usable and meaningful way. At the same time, broader integration, particularly in the area of active dictation and voice-enabled interaction within the ordinary desktop translation workflow, remains incomplete. The Romanian case therefore points to a real and growing institutional presence of ASR within the European institutions, and more specifically



within the European Parliament, together with promising practical value. At the same time, however, it also points to an ongoing process of development and implementation that still depends on further institutional support, approved solutions, and more systematic quality evaluation.

## **5.2 Strengths and limitations**

Having discussed the main findings and their broader implications, it is also important to reflect critically on the contribution and scope of the present study. This subchapter therefore considers both the main strengths of the thesis and the limitations that shape the interpretation of its results. Such a reflection helps to position the study more clearly within the current discussion on ASR in European institutional settings and to indicate more precisely what this thesis was able to contribute, as well as where its scope necessarily remained limited.

One of the main strengths of this thesis lies in its choice of topic and perspective. The study addresses an area that is increasingly relevant in professional language work and in European institutional communication, yet still insufficiently documented from the point of view of concrete institutional workflows and specific languages such as Romanian. By focusing not only on ASR as a technical phenomenon but also on its relevance, usability and acceptance in a multilingual European institutional environment, the thesis contributes to a more practice-oriented understanding of how speech technologies may enter professional communication, translation and documentation settings. This user-centred and workflow-centred perspective is, in itself, an important strength, because it allows the discussion to move beyond abstract claims about technological progress.

A second strength is the institutional embeddedness of the study. The thesis does not examine Romanian ASR in a generic environment, but in the concrete setting of the UE institutions and more precisely, in the European Parliament. This gives the research particular value because European institutions are highly multilingual, strongly regulated and operationally demanding environments in which the adoption of language technologies is shaped not only by efficiency but also by accessibility, confidentiality, risk management and governance. Studying ASR in such a context makes it possible to show more clearly which aspects of the technology are institutionally meaningful and which remain limited by practical conditions. The present thesis therefore contributes evidence from a professional setting that is especially relevant for European Translation Studies, institutional communication and applied language technology.

A third strength concerns the Romanian-language focus. Much of the existing ASR literature remains concentrated on larger languages or on technical evaluation settings that do not say much about user experience in institutional work. By centring Romanian, the thesis responds to a clear gap in the literature. It shows that Romanian should be discussed as a language that is already present in European institutional ASR systems and that deserves analysis in relation to real workflows and professional perceptions. This contribution helps position Romanian not at the margins of the ASR discussion, but within the broader European multilingual context.

The mixed-methods design is another important strength of the thesis. By combining two interviews, a presentation, a short try-out and a questionnaire, the study constructs a layered

empirical picture of ASR in the Romanian-language institutional context. The expert interview provides insight into institutional development, evaluation practices and operational constraints; the Verba interview adds a concrete user perspective on everyday transcription work; the presentation and try-out extend the discussion towards possible workstation-level uses; and the questionnaire captures the views of potential users after a structured introduction to the demonstrated tools. These methods do not produce the same kind of evidence, but this is precisely one of their main advantages.

Together, they make it possible to identify a particularly important pattern: ASR may already be operational and meaningful in one institutional workflow while remaining practically unavailable in another; users may express openness towards the technology even where institutional conditions prevent routine access to it; and Romanian may appear usable in certain settings while still lacking more systematic language-specific evaluation. Insights of this kind emerge precisely because the thesis does not rely on a single method or perspective. Instead, it approaches ASR from several complementary angles and thus offers a more differentiated understanding of its institutional role.

A further strength of the thesis lies in its clear and consistent approach. The study applies the same three criteria throughout, namely relevance, usability and acceptance, which creates a strong link between the theoretical foundation, the methodology and the analysis of the findings. This also makes comparison possible across the different forms of ASR discussed in the thesis. Although live transcription, asynchronous transcription and active dictation or voice interaction serve different purposes and operate under different institutional conditions, they can all be examined through the same guiding questions: Is the technology relevant for the task? Is it usable in practice? Is it acceptable within the institutional environment?

At the same time, the thesis remains closely grounded in practice by showing concretely how institutional constraints shape the actual possibilities of ASR use. The blocked Windows voice typing function, the linguistic and licensing limits of Dragon, the live-only nature of the Speech-to-Text platform and the controlled role of Verba all illustrate how governance, confidentiality and workflow compatibility influence implementation in real settings. This combination of analytical clarity and practical detail is an important strength of the study because it also increases the practical relevance of the thesis for future discussions on Romanian ASR and CAT-tool integration, while complementing the still limited public information available in EU institutional reporting.

At the same time, the thesis has clear limitations. The first limitation is scale. The study is exploratory and limited in scope, since it focuses on one institution, one language unit and a relatively small number of participants. The findings therefore cannot be generalised automatically to all Romanian-language professionals in European institutions, still less to all institutional settings or all professional translation environments. The value of the study lies rather in the depth and specificity of the case than in broad representativeness.

A second limitation concerns the empirical material itself. On the interview side, the live ASR findings are based on one expert interview with the Head of the Speech-to-Text Unit, while the Verba findings rely on one short interview with one experienced Romanian proofreader. These interviews are highly informative, but they cannot provide a complete

picture of user experience or institutional practice across different languages, EP teams and user profiles.

Likewise, the questionnaire offers only a partial institutional snapshot: it was completed by nine respondents from a unit of twenty-two staff members and captures reactions after one presentation and, in some cases, one brief try-out. For this reason, the questionnaire documents perceptions, expectations and conditional openness more than long-term user behaviour. Related to this is the limitation of self-reporting more generally. The thesis cannot claim to measure actual productivity gains, real correction times or longer-term adoption patterns in Romanian-language translation work. It captures informed attitudes and expectations rather than performance in action.

A third limitation lies in the absence of quantitative performance metrics. The thesis discusses usability in terms of perceived correction effort, manageability and workflow fit, but it does not measure word error rate, latency, exact time savings or detailed Romanian-specific error frequencies. This applies both to the live platform and to Verba, though for different reasons. In the case of the live platform, access to persistent output is limited by the live-only nature of the service and by data-protection requirements for post accessing transcripts. In the case of Verba, the study had access to user experience, but not to the system's internal architecture or to systematic performance data. As a result, the thesis remains stronger on the qualitative and interpretive side than on the technical-evaluative one.

A further limitation concerns the workstation-level ASR solutions examined in the presentation. The range of tools was necessarily limited and does not represent a complete overview of all possible solutions for active ASR use in the European Parliament. Windows voice typing, Dragon and Trados Online Editor were selected because they were the most visible, relevant and practically demonstrable examples for the Romanian institutional translation context. In particular, Trados Online Editor was included not because it reflects the regular EP desktop workflow, but because no other equally practical and demonstrable solution for Romanian ASR integrated into a CAT environment was available for the purposes of this study. Its role in the thesis is therefore illustrative and indicative of possible future development rather than fully operational within the current EP workflow.

The thesis is also limited by its time frame and by the broad scope of the topic. It reflects the situation of ASR in 2025-2026, in a field that is developing quickly. This means that ASR tool availability, institutional rules and Romanian-language support may change over time. This is particularly relevant in the Romanian context, where both language resources and institutional use are still developing. Another limitation is that the thesis covers several areas at once, namely accessibility, transcription, translation workflow and Romanian-language ASR. This broad perspective is comprehensive, but it also means that not every area could be examined in the same level of detail. A narrower study could have focused more deeply on one tool, one type of ASR use or one workflow. The present thesis instead aimed to provide a broader institutional overview.

These limitations help define more clearly the kind of contribution this thesis makes. It is not intended as a definitive benchmarking study, a large-scale survey or a controlled productivity experiment. Rather, it is an exploratory institutional case study that examines how ASR currently appears, functions and is perceived in one concrete Romanian-language setting

within the European Parliament. Within this scope, the thesis makes a specific contribution that has so far remained limited in the literature: it offers a Romanian-focused, workflow-sensitive and institutionally grounded account of ASR across different forms of use.

In this sense, the study contributes most clearly at the intersection of Romanian, institutional practice and professional user perception. It shows that Romanian ASR in European institutions should not be discussed only in terms of language coverage, data scarcity or technical availability, but also in terms of actual use, correction effort, institutional legitimacy and user readiness. It also shows that the main question for future development is no longer whether ASR exists in the institution, because it already does in several forms, but how these forms can be aligned more effectively with specific workflows, professional needs and user groups.

The strengths and limitations of the thesis therefore point in the same direction. Its main strength lies in showing ASR as an institutional practice shaped by function, language and governance; its main limitations lie in the restricted scale of the case, the exploratory character of the evidence and the absence of broader quantitative evaluation. At the same time, precisely because the study makes these limits visible, it also helps define the next steps for future research and implementation. In this respect, the exploratory character of the thesis is not a weakness in itself, but a reflection of the current stage of Romanian institutional ASR: a field that is already present and relevant, but still developing, still unevenly integrated and still in need of closer research.

### **5.3 Practical implications and recommendations**

Building on the findings and limitations discussed in the previous subchapters, this section considers what they could imply both for future ASR implementation in European institutions and for further research in this field. In this way, the subchapter identifies possible directions not only for future integration strategies in EU institutions in general and, more specifically, in the European Parliament and the Romanian context, but also for further research on ASR under institutional and multilingual conditions.

A first practical implication of the thesis is that future ASR implementation projects, as well as future research in European institutions, should be planned in relation to specific workflows rather than approached through one general model of technological uptake. The present study suggests that live accessibility services, internal transcription workflows and workstation-level use serve different purposes, operate under different constraints and involve different thresholds of acceptance. In the broader EU setting, including the European Parliament, future implementation strategies would therefore benefit from distinguishing more clearly between these forms of use.

Accessibility-oriented ASR, for example, could continue to be developed primarily for real-time access and multilingual communication and be evaluated accordingly, since its main institutional value lies above all in accessibility and immediate communication needs rather than in perfect textual precision. Internal transcription tools, by contrast, could be further improved and evaluated as human-revisable workflow support, because their output is linked to institutional documentation and publication and therefore requires stronger post-correction standards. Seen in this way, the main recommendation is that future institutional development

and research should not treat ASR as a single general solution but should align each form of ASR more closely with the concrete workflow, quality requirements and institutional purpose for which it is intended.

Workstation-level ASR, however, still requires further investigation, since the present thesis was able to examine only a limited number of currently available or demonstrable solutions. This follows directly from one of the main limitations of the study, namely that active ASR use in the ordinary Romanian-language desktop translation workflow remains more conceivable than fully implemented under current EP conditions. For this reason, future institutional implementation in this area would benefit from a gradual and exploratory approach.

In this sense, one possible direction would be to examine whether a commercially available ASR solution could be adapted and institutionally governed in such a way that it supports Romanian-language dictation and voice interaction on EP laptops without compromising confidentiality or data-security requirements. Another possible direction would be to explore whether the technical experience, workflow knowledge and language resources already accumulated in other institutional ASR contexts could support future ASR translation-related development under appropriate legal and technical conditions. Because this area remains less developed, active ASR would be better tested first in small, task-specific pilots than through large-scale deployment. A useful next step could therefore be a limited pilot involving lower-risk tasks such as note-taking or email drafting, followed only later by more demanding CAT-tool scenarios if the first results prove workable.

Another important recommendation concerns terminology and language-specific adaptation. The thesis has shown repeatedly that the usability of ASR in institutional settings depends not only on general language coverage, but also on how well the system handles names, acronyms, committee titles, programme names and other recurrent forms of parliamentary language. In the Romanian case, this appears especially relevant, since the findings suggest that certain difficulties are linked less to Romanian recognition in general than to institutional vocabulary, lexical updating and domain-specific variation. A practical step could therefore be the introduction of more regular terminology update cycles for ASR-related tools, for example after significant changes in institutional nomenclature or in connection with new legislative priorities and public programmes. In research terms, this also suggests the usefulness of studying Romanian ASR not only as a general-language problem, but as a domain-specific one shaped by parliamentary discourse and institutional phraseology.

The study also points to the value of structured user feedback. One important lesson of both the user interview and the questionnaire is that users can often identify recurring problems very clearly, even if they do not have access to the technical side of the system. Future institutional development could therefore benefit from a simple feedback mechanism through which users can report recurrent lexical problems, recognition errors or workflow difficulties. In practice, this might take the form of a short internal reporting template, a shared terminology note or a more regular exchange between language professionals and those responsible for updating the system with DG TRAD of the European Parliament, or in general in EU institutions. Such a structure would be useful not only for improvement in practice, but also for future research, because it would gradually create a stronger empirical basis for more detailed Romanian-specific evaluation and for clearer comparison across tools, workflows and institutional contexts.

A further recommendation concerns the role of human oversight. The findings of the thesis do not suggest that ASR is moving towards the replacement of human linguistic work in European institutions. In the case of internal transcription, the need for human revision remains essential, especially in politically sensitive parliamentary discourse. The interviews and institutional context discussed in the thesis make clear that, in workflows such as those supported by Verba, revision is not a secondary stage but a necessary condition of acceptability, precisely because the final text is intended for official publication and must not misrepresent what a Member intended to say. In the case of the live Speech-to-Text platform, the role of ASR is different again. The interview with the Head of the Speech-to-Text Unit made clear that the platform, even when it includes live translation, is not intended to replace human interpreters. Its main purpose is to provide accessibility and practical multilingual support in a setting where several events may take place simultaneously and where immediate written access may be useful, especially in less politically sensitive contexts.

In practical terms, future implementation strategies should therefore continue to treat human revision, correction and linguistic mediation not as signs of technological failure, but as part of a realistic and professionally appropriate use of ASR in institutional settings. This also has implications for research, since future studies may benefit from focusing not only on recognition output itself, but also on the relationship between output quality, correction effort, risk level and workflow usefulness.

In relation to CAT-tool use and translation-related tasks, the findings suggest that future experimentation may be more promising if it begins with hybrid forms of support rather than with full dictation from scratch. Both the theoretical discussion and the user-oriented findings indicate that institutional translation work often consists largely of revision, post-editing and navigation within segmented environments. For this reason, future development could usefully explore speech-supported revision, selective voice commands, punctuation insertion, navigation between segments and voice-assisted correction of pretranslated text. A practical next step might therefore be a small comparative pilot in which Romanian-language professional perform a short revision or post-editing task of pre-translated segment in Trados Studio, once in the usual way and once with voice support, so that correction effort, user perception and practical usefulness can be compared more directly. Such a design would also address one of the limitations of the present thesis, namely the absence of longer controlled practical testing and more systematic task-based comparison.

Training and awareness also emerge as practical priorities. The questionnaire findings suggest that even respondents with little previous experience of ASR reacted more positively once the technology had been explained and demonstrated in a structured way. This indicates that future institutional implementation may depend not only on better implemented tools, but also on better preparation of users. Short, practice-oriented training could help users understand what ASR can realistically do, where it may be useful, where it may not be appropriate, and how institutional rules on confidentiality and data handling apply. In the Romanian context, where the questionnaire suggests conditional openness rather than resistance, this kind of training may be especially important for turning general interest into informed and sustainable professional use.

The limitations identified in the thesis also point quite clearly to what further research could usefully do. Since the present study was exploratory, limited in scale and stronger on the

qualitative side than on the quantitative one, one important next step would be a larger and more systematic study involving more participants, possibly across several language units or even across several EU institutions. Such a study could combine interviews, repeated observation and follow-up questionnaires over a longer period, making it possible to examine not only first reactions, but also more sustained patterns of ASR use.

A second important direction would be direct performance measurement. Future research could usefully compare typing, post-editing and speech-supported work in short, controlled tasks, while also examining correction time, error types and user effort more systematically. In the Romanian case, a further valuable step would be the development of a more detailed error typology based on institutional material, distinguishing more clearly between terminology-related problems, punctuation, diacritics, homophonous forms, disfluency and speaker variability. This would respond directly to one of the main limitations of the present thesis and would provide a stronger basis for future implementation decisions.

Finally, future research could also broaden the range of perspectives included. The present thesis focused mainly on expert and professional-language users, which was appropriate for the research design, but the broader institutional role of ASR also includes external users of accessibility-oriented services. In the wider EU setting, it may therefore be useful to complement the present findings with user-focused research on how live multilingual ASR is experienced by those who depend directly on it. In practical terms, this could involve short post-event questionnaires, targeted accessibility feedback, or small studies comparing user satisfaction across event types and languages. Such work would help connect institutional design goals more closely with actual user needs and would provide a fuller picture of ASR in European institutional communication.

Taken together, the practical implications and research directions identified in this thesis suggest that the future of ASR in European institutions is likely to depend on gradual, secure and workflow-sensitive integration. The study does not point towards one single model of implementation, but rather towards several connected paths: clearer differentiation between workflows, stronger secure infrastructure, more systematic terminology updating, structured user feedback, selective workstation-level pilots, practice-oriented training and more detailed future research.

At the same time, the findings also suggest that implementation in EU institutions is structurally demanding. These institutions function as complex organisational systems in which new tools cannot simply be introduced on the basis of technical availability alone. Deployment typically depends on prior authorisation, coordination across several services, compliance with data-protection and security requirements, technical support, access management and, in some cases, licensing or procurement decisions. For this reason, future ASR development in the broader EU setting, including the European Parliament, is likely to require not only technological progress, but also sustained institutional coordination across units with different responsibilities.

In this context, ASR may become increasingly important not as a uniform technological solution, but as a flexible set of tools whose value depends on how well they are aligned with institutional responsibility, multilingual realities and real professional use. In the Romanian context, this future appears promising, but it remains closely linked to continued institutional support, clearer evaluation and more detailed language-specific follow-up.

## 5.4 Final conclusion

In a world in which speech technologies have become an increasingly familiar presence in everyday life, this thesis set out to examine what happens when automatic speech recognition is observed not as a general technological trend, but in a concrete professional and institutional setting. More specifically, it investigated the relevance, usability and acceptance of Automatic Speech Recognition (ASR) in European institutions through a Romanian-language case study at the European Parliament. What emerges from this case study is that ASR is no longer simply a future possibility in multilingual institutional communication. It is already present, already meaningful and already capable of supporting certain institutional needs. At the same time, its value is neither uniform nor automatic. It depends on the form in which ASR is used, the workflow into which it is introduced, the language conditions under which it operates and the institutional environment that enables or restricts its use.

From this perspective, the main conclusion of the thesis is not simply that ASR is present in the European Parliament, but that its institutional meaning depends on how it is situated. Where immediate access to spoken content matters most, ASR can already support accessibility and multilingual communication in meaningful ways. Where institutional documentation is concerned, it can reduce effort while remaining dependent on human revision and professional control. Where translation-related desk work is concerned, its role remains more tentative, not because the technology is irrelevant in principle, but because institutional availability, workflow fit and secure implementation have not yet fully converged.

The Romanian case gives this broader conclusion its clearest focus. This thesis places Romanian not merely among the languages formally covered by institutional ASR, but within the actual realities of institutional practice. In doing so, it shows that Romanian should not be discussed only in terms of technical support or formal inclusion. The more important question is how Romanian functions when ASR meets parliamentary discourse, institutional terminology, correction effort and professional expectations. In this sense, the thesis shifts the discussion from the fact of inclusion to the quality of integration. Romanian is already part of the institutional ASR landscape, but its role within that landscape still depends on how carefully technological development is connected with real linguistic work.

The study also makes clear that the future of ASR in European institutions will not be shaped by technological progress alone. European institutions are complex organisational environments in which implementation depends on authorisation, coordination between different services, data-protection requirements, IT restrictions and different degrees of communicative sensitivity. For this reason, ASR is unlikely to advance through one simple moment of adoption. Its development is more likely to take the form of gradual institutional alignment, in which technical possibility, governance and professional trust have to move forward together. In such a setting, human expertise remains central. The findings do not point towards the disappearance of interpreters, translators or proofreaders, but towards new forms of cooperation in which ASR can support their work.

Seen in this light, the title of the thesis can also be read as its final answer. The relevance of ASR in European institutions lies in the fact that it already responds to concrete institutional needs. Its usability is reflected in the practical value it can already offer in certain workflows. Its acceptance can be seen in the openness that professionals already show towards it under



appropriate conditions. Yet all three remain dependent on fit: fit with the workflow, fit with the language, fit with institutional rules and fit with professional responsibility. In the Romanian case, the study therefore points not to a marginal or speculative future, but to a credible one: ASR is no longer only a possibility on the horizon, but not yet a finished institutional routine. It is a developing practice whose next stage will depend on whether institutions are willing to support it with the same seriousness with which they already depend on multilingual communication itself.

## 6. Bibliography

- Arora, Shipra J. & Singh, Rishi Pal (2012). Automatic speech recognition: a review. *International Journal of Computer Applications* 60(9), 34–44. <https://doi.org/10.5120/9722-4190>
- Avram, Andrei Marius, Păiș, Vasile & Tufiș, Dan (2020). Towards a Romanian end-to-end automatic speech recognition based on DeepSpeech2. *Proceedings of the Romanian Academy, Series A* 21(4), 395–402.
- Ciobanu, Dragoș & Secară, Alina (2019). Speech recognition and synthesis technologies in the translation workflow. In: O'Hagan, Minako (ed.) *The Routledge Handbook of Translation and Technology*. Abingdon/New York: Routledge, 91–106.
- Di Nuovo, Elisa (2023). An experiment in error analysis of real-time speech machine translation using the example of the European Parliament's Innovation Partnership. In: Boschetti, Federico; Lebani, Gianluca E.; Magnini, Bernardo & Novielli, Nicole (eds.) *Proceedings of the Ninth Italian Conference on Computational Linguistics (CLiC-it 2023)*. Venice: CEUR Workshop Proceedings 3596, 185–195.
- Dresing, Thorsten & Pehl, Thorsten (2015). *Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende*. Marburg: dr. dresing & pehl GmbH.
- Europarl-ST (n.d.). Europarl-ST corpus (release v1.1). <https://www.ml1p.upv.es/euoparl-st/> (last accessed 09.04.2026).
- European Commission (2024). Automated Speech Recognition. Shaping Europe's Digital Future. Updated 16 May 2025. <https://digital-strategy.ec.europa.eu/en/news/automated-speech-recognition> (last accessed 06.01.2026).
- European Commission (n.d.a). Digital Europe - AI-based Multilingual Services. <https://language-tools.ec.europa.eu/> (last accessed 09.04.2026).
- European Commission (n.d.b). AI-based multilingual services – using EU language data to innovate. [https://translation.ec.europa.eu/language-data-and-ai-using-ai-break-down-language-barriers/ai-based-multilingual-services-using-eu-language-data-innovate\\_en](https://translation.ec.europa.eu/language-data-and-ai-using-ai-break-down-language-barriers/ai-based-multilingual-services-using-eu-language-data-innovate_en) (last accessed 09.04.2026).
- European Parliament (2008). Guide to the obligations of officials and other servants of the European Parliament: Code of conduct, adopted by the Bureau on 7 July 2008. [https://www.euoparl.europa.eu/RegData/PDF/406411\\_EN.pdf](https://www.euoparl.europa.eu/RegData/PDF/406411_EN.pdf) (last accessed 08.01.2026).

European Parliament (2010). Rapport annuel d'activités. <https://www.europarl.europa.eu/document/activities/cont/201008/20100825ATT80370/20100825ATT80370FR.pdf> (last accessed 09.04.2026).

European Parliament (2025). Petition No 0886/2025 by I. L. P. (Romanian) on regulation of competitions for animal traction in the European Union. <https://www.europarl.europa.eu/petitions/en/petition/content/0886%252F2025/html/Petition-No-0886%252F2025-by-I.-L.-P.-%2528Romanian%2529-on-regulation-of-competitions-for-animal-traction-in-the-European-Union> (last accessed 09.04.2026).

European Parliament (2026a). European Central Bank – annual report 2025 (debate). Verbatim report of proceedings, 9 February 2026. [https://www.europarl.europa.eu/doceo/document/CRE-10-2026-02-09-ITM-015\\_EN.html](https://www.europarl.europa.eu/doceo/document/CRE-10-2026-02-09-ITM-015_EN.html) (last accessed 12.04.2026).

European Parliament (2026b). Sitting of 09-02-2026: video-on-demand entry for Vlad Vasile-Voiculescu. <https://www.europarl.europa.eu/plenary/en/vod.html?mode=unit&vodLanguage=EN&playerStartTime=20260209-20:15:55&playerEndTime=20260209-20:17:31> (last accessed 12.04.2026).

European Parliament (n.d.a). Debates and Videos. Plenary. <https://www.europarl.europa.eu/plenary/en/debates-video.html> (last accessed 12.04.2026).

European Parliament (n.d.b). DG TRAD Help page. <https://www.europarl.europa.eu/trad/telework/> (last accessed 13.04.2026).

European Parliament, Directorate-General for Translation (2024). *Annual Activity Report 2023*. [https://www.europarl.europa.eu/cmsdata/285562/9\\_TRAD%20-%20AAR%202023.pdf](https://www.europarl.europa.eu/cmsdata/285562/9_TRAD%20-%20AAR%202023.pdf) (last accessed 10.01.2026).

Furui, Sadaoki (2005). 50 years of progress in speech and speaker recognition research. *ECTI Transactions on Computer and Information Technology* 1(2), 64–74. <https://doi.org/10.37936/ecti-cit.200512.51834>

Georgescu, Alexandru-Lucian, Cucu, Horia, Buzo, Andi & Burileanu, Corneliu (2020). RSC: a Romanian read speech corpus for automatic speech recognition. In: *Proceedings of the Twelfth Language Resources and Evaluation Conference (LREC 2020)*. Marseille: European Language Resources Association, 6606–6612.

Juang, Biing-Hwang & Rabiner, Lawrence R. (2004). Automatic speech recognition – a brief history of the technology development. PDF manuscript, UC Santa Barbara.

[https://web.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/354\\_LALI-ASRHistory-final-10-8.pdf](https://web.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/354_LALI-ASRHistory-final-10-8.pdf) (last accessed 09.04.2026).

Kheddar, Hamza, Hemis, Mustapha & Himeur, Yassine (2024). Automatic speech recognition using advanced deep learning approaches: a survey. *Information Fusion* 109, 102422.

LDS (2025). LDS Workshop on Speech Recognition Solutions. [https://language-data-space.ec.europa.eu/events/lds-workshop-speech-recognition-solutions-2025-07-01\\_en](https://language-data-space.ec.europa.eu/events/lds-workshop-speech-recognition-solutions-2025-07-01_en) (last accessed 10.01.2026).

LDS (n.d.). European Language Data Space. [https://language-data-space.ec.europa.eu/index\\_en](https://language-data-space.ec.europa.eu/index_en) (last accessed 10.01.2026).

LINAGORA AI (n.d.). LinTO.ai, sovereign Voice AI for your applications. <https://www.linagora.ai/en/solutions/linto> (last accessed 12.04.2026).

Mavrič, Valter (2022). Accessibility for the deaf and hard of hearing through language technology: real-time speech-to-text and machine translation tool of the European Parliament. Keynote slides, Translating and the Computer 44. [https://asling.org/tc44/slides/TC44-Keynote-Mavric\\_Valter-Accessibility%20for%20the%20deaf%20and%20hard%20of%20hearing.pdf](https://asling.org/tc44/slides/TC44-Keynote-Mavric_Valter-Accessibility%20for%20the%20deaf%20and%20hard%20of%20hearing.pdf) (last accessed 12.04.2026).

Mayring, Philipp (2022). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim/Basel: Beltz.

Microsoft (2025). Azure Translator in Foundry Tools frequently asked questions. <https://learn.microsoft.com/en-us/azure/ai-services/translator/faq> (last accessed 09.04.2026).

Microsoft (n.d.). Azure Translator in Foundry Tools documentation. <https://learn.microsoft.com/en-us/azure/ai-services/translator/> (last accessed 12.04.2026).

Microsoft Windows (n.d.). Use voice typing to talk instead of type on your PC. <https://support.microsoft.com/en-us/windows/use-voice-typing-to-talk-instead-of-type-on-your-pc-fec94565-c4bd-329d-e59a-af033fa5689f> (last accessed 09.01.2026).

Mozilla Common Voice (2026). Common Voice Scripted Speech 25.0 - Romanian [Data set]. <https://mozilladatasetcollective.com/datasets/cmn2e8rmi0116mm07vxurptse> (last accessed 09.04.2026).

Nuance (2023). What languages are available for Dragon 15 Desktop Products? [https://nuance.custhelp.com/app/answers/detail/a\\_id/6280/~/\\_what-languages-are-available-for-dragon-15-desktop-products%3F](https://nuance.custhelp.com/app/answers/detail/a_id/6280/~/_what-languages-are-available-for-dragon-15-desktop-products%3F) (last accessed 19.11.2025).

- RWS (n.d.). Dictation in Trados Online Editor. <https://docs.rws.com/trados-enterprise-%26-accelerate-791611/dictation-in-trados-online-editor-1123284> (last accessed 01.11.2025).
- Saldanha, Gabriela & O'Brien, Sharon (2013). *Research Methodologies in Translation Studies*. London/New York: Routledge.
- Wang, Changan, Rivière, Morgane, Lee, Ann, Wu, Anne, Talnikar, Chaitanya, Haziza, Daniel, Williamson, Mary, Pino, Juan & Dupoux, Emmanuel (2021). VoxPopuli: a large-scale multilingual speech corpus for representation learning, semi-supervised learning and interpretation. In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Online: Association for Computational Linguistics, 993–1003.
- Yu, Dong & Deng, Li (2015). *Automatic Speech Recognition: A Deep Learning Approach*. London: Springer.
- Zapata, Julián & Søeborg Kirkedal, Andreas (2015). Assessing the performance of automatic speech recognition systems when used by native and non-native speakers of three major languages in dictation workflows. In: Megyesi, Beáta (ed.) *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)*. Vilnius: Linköping University Electronic Press, 201–210.
- Zapata, Julián, Castilho, Sheila & Moorkens, Joss (2017). Translation dictation vs. post-editing with cloud-based voice recognition: a pilot experiment. In: *Proceedings of Machine Translation Summit XVI: Commercial MT Users and Translators Track*. Nagoya, Japan, 116–129.
- Zapata, Julián, Secară, Alina & Ciobanu, Dragoş (2023). Past, present and future of speech technologies in translation — life beyond the keyboard. In: Moorkens, Joss & Sosoni, Vilelmini (eds.) *Translating and the Computer 44 (TC44): Proceedings*. Geneva: Editions Tradulex, 16–26.

## 7. Appendices

### Appendix A - Interview with the Head of EP STT Unit

#### Interview questions

| No. | English                                                                                                                                                                                | Romanian                                                                                                                                                                                                      |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|     | <b>General Questions</b>                                                                                                                                                               | <b>Întrebări Generale</b>                                                                                                                                                                                     |
| 1   | Could you describe your key responsibilities as Head of the Speech-to-Text Unit?                                                                                                       | Puteți descrie principalele responsabilități pe care le aveți în calitate de Șef al Unității Speech-to-Text?                                                                                                  |
| 2   | What are the main objectives and tasks of your Unit within the European Parliament?                                                                                                    | Care sunt obiectivele și sarcinile esențiale ale acestei unități în cadrul Parlamentului European?                                                                                                            |
| 3   | Can you describe how ASR technologies are currently being used or developed in your Unit?                                                                                              | Cum sunt implementate sau dezvoltate în prezent tehnologiile ASR în cadrul unității dumneavoastră?                                                                                                            |
| 4   | In your opinion, what are the main advantages of using ASR tools in an institutional context such as the European Parliament?                                                          | În opinia dumneavoastră, care sunt principalele avantaje ale utilizării instrumentelor ASR într-un context instituțional, precum Parlamentul European?                                                        |
| 5   | What challenges or limitations have you encountered when developing or implementing ASR systems within the European Parliament?                                                        | Ce provocări sau limitări ați întâmpinat în procesul de dezvoltare sau implementare a tehnologiilor ASR în cadrul Parlamentului European?                                                                     |
|     | <b>Questions on the internally developed Speech-to-Text platform</b>                                                                                                                   | <b>Întrebări legate de platforma Speech-to-Text recent dezvoltată</b>                                                                                                                                         |
| 6   | Could you describe the development process of the Speech-to-Text platform within the European Parliament?                                                                              | Puteți descrie procesul de dezvoltare al platformei Speech-to-Text din cadrul Parlamentului European?                                                                                                         |
| 7   | Could you elaborate on how the platform has been adapted to process parliamentary discourse, particularly in terms of terminology?                                                     | Puteți oferi mai multe detalii despre modul în care platforma a fost adaptată pentru a prelucra discursurile parlamentare, în special din punct de vedere terminologic?                                       |
| 8   | Have there been any benchmarking experiments comparing the workflow with ASR for automatic live transcription to the workflow with a respeaker using ASR actively for live captioning? | Au fost realizate experimente pentru a compara fluxul de lucru care utilizează ASR pentru transcriere automată în direct cu cel care implică un respeaker ce folosește activ ASR pentru captioning în direct? |

|    |                                                                                                                                                                                                   |                                                                                                                                                      |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
|    | Which languages were included and what were the main findings?<br><br>If no such benchmarking has been conducted, could you explain why?                                                          | Ce limbi au fost incluse și care au fost principalele concluzii?<br><br>Dacă nu s-au efectuat astfel de teste, care au fost motivele?                |
|    | <b>Romanian-focused questions</b>                                                                                                                                                                 | <b>Întrebări referitoare la limba română</b>                                                                                                         |
| 9  | How well is Romanian currently supported by the Speech-to-Text platform?                                                                                                                          | Cât de bine funcționează recunoașterea automată a vorbirii pentru limba română pe platforma Speech-to-Text?                                          |
| 10 | Are there any linguistic / phonetic challenges when applying ASR technology to Romanian, compared to other EU languages?                                                                          | Există provocări lingvistice / fonetice specifice atunci când se aplică tehnologia ASR la limba română, în comparație cu alte limbi oficiale ale UE? |
|    | <b>Final Question</b>                                                                                                                                                                             | <b>Întrebare finală</b>                                                                                                                              |
| 11 | If there were no budgetary or technical limitations, what improvements or innovations would you prioritize to advance the development and use of ASR technologies within the European Parliament? | Ce ați schimba sau îmbunătăți în domeniul ASR din cadrul Parlamentului European, dacă nu ar exista deloc constrângeri tehnice sau financiare?        |

## Transcript and recordings

### Recording files (RO):

<https://drive.google.com/file/d/1VxyKC7H-akgSGD2d3QEr2WbsykfsjiI/view?usp=sharing>

<https://drive.google.com/file/d/16KuDE38bSzKRjag6sQ61Obo6cUvblvdF/view?usp=sharing>

| English Translation                                                                                                                                                                                                                  | Romanian Original                                                                                                                                                                                                                                                        |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Interviewer:</b> I would like to ask you, in general, what the main responsibilities are in your capacity as head of the unit, or, more generally, what the unit's tasks and objectives are.                                      | <b>Intervievator:</b> Aș vrea să te întreb, în general, care sunt responsabilitățile principale în calitate de șef al unității sau, în general, task-urile și obiectivele unității?                                                                                      |
| <b>Interviewee:</b> So, the unit's task and objective are the development, implementation and provision of an application or a series of applications that ensure automatic speech recognition, transcription and translation in all | <b>Intervievat:</b> Deci, task-ul și obiectivul unității sunt dezvoltarea, implementarea și asigurarea unei aplicații sau a unei serii de aplicații care să asigure recunoașterea automată a vorbirii, transcrierea și traducerea în toate cele 24 de limbi oficiale ale |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 24 official languages of the European Union, primarily for accessibility purposes for persons with disabilities and, secondarily, for practical reasons for Members of the European Parliament, staff and their accredited assistants, that is, MEPs and staff.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Uniunii Europene, în principal în scopul accesibilității pentru persoanele cu dizabilități și, în scop secundar, din motive practice pentru europarlamentari, personal și asistenții lor acreditați, adică MEPs and staff.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Interviewer:</b> I see. But, in connection with translation, in that case, is the aim greater efficiency, in place of interpreters, live?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | <b>Interviewator:</b> Am înțeles. Dar, în legătură cu traducerea, în acel caz, se țintește spre o mai bună eficiență, în locul interpreților, în direct?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| <b>Interviewee:</b> No, no – and, officially, we do not aim to replace interpreters. Our main goal is accessibility and providing practical, instantaneous, live translation and transcription.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | <b>Interviewat:</b> Nu, nu – și, zicem oficial, nu urmărim să înlocuim interpreții. Scopul nostru principal este accesibilitatea și oferirea unei traduceri și transcrieri practice, instantanee, în direct.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <b>Interviewer:</b> You mentioned these ASR technologies as the unit's main objectives. Could you tell me how these automatic speech recognition technologies are currently implemented within your unit?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | <b>Interviewator:</b> Ați menționat aceste tehnologii ASR ca obiective principale ale unității. Ați putea să-mi spuneți cum sunt implementate în momentul de față aceste tehnologii de recunoaștere automată a vorbirii în cadrul unității dumneavoastră?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Interviewee:</b> We have a platform made up of several modules. The platform is the one developed for the European Commission, the Linto Platform. After that, we use a number of linguistic modules: a language identification module, then a transcription module and, after that, a translation module, for which, for the moment, we use Microsoft Azure.                                                                                                                                                                                                                                                                                                                                                                                                                                                    | <b>Interviewat:</b> Avem o platformă din mai multe module. Platforma este cea dezvoltată pentru Comisia Europeană, Linto Platform. După aceea folosim niște module lingvistice: un modul de identificare a limbii, apoi un modul de transcriere și, după aceea, un modul de traducere, pentru care, pe moment, folosim Microsoft Azure.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <b>Interviewer:</b> So, at the moment, are you using a commercial resource?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | <b>Interviewator:</b> Și anume, în momentul de față, folosiți o resursă comercială?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <p><b>Interviewee:</b> Yes. But what is important to note is that, in the European Parliament, a commercial source cannot be used off the shelf. The first reason is that the language used on the street and picked up by AI is very different from the language spoken in Parliament. For example, if we look at it quantitatively, Romanian-language usage is very heavily influenced by manele culture. If you do a search, the AI goes off and searches on its own, on YouTube, on TikTok. And that type of language is manele-influenced and does not correspond to what is spoken in Parliament.</p> <p>The second reason is the structure of the language itself – its grammar and language use. It is different from what AI let loose on the internet means (AI left to roam freely on the internet).</p> | <p><b>Interviewat:</b> Da. Dar ce este important de notat este că, în Parlamentul European, o sursă comercială nu poate fi folosită de-a gata (off the shelf). Primul motiv este că limbajul folosit pe stradă și preluat de AI este foarte diferit de limba vorbită în Parlament. De exemplu, dacă luăm cantitativ, limbajul în română este foarte manelizat. Dacă dai search, AI-ul se duce și caută singur, pe YouTube, pe TikTok. Iar acel tip de limbaj este unul manelizat și nu corespunde cu ce se vorbește în Parlament.</p> <p>Al doilea motiv este structura limbii în sine – a gramaticii, a limbajului. Este diferită de ceea ce înseamnă AI let loose on the internet (AI-ul lăsat de capul lui pe internet).</p> <p>Al treilea motiv este diferența dintre limba vorbită în interiorul Uniunii Europene și cea vorbită în afara Uniunii (în cazul limbii române nu se aplică, nefiind o limbă vorbită intercontinentală). AI-ul nu este la</p> |



|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>The third reason is the difference between the language spoken within the European Union and the one spoken outside the Union (in the case of Romanian this does not apply, since it is not an intercontinentally spoken language). AI is not yet at the level of making that distinction. For example, there are differences between French as spoken in France and French as spoken in Quebec or Senegal, between Portuguese as spoken in Portugal and Portuguese as spoken in Brazil, or between Spanish as spoken in Spain and Spanish as spoken in Mexico or Argentina. From what I understand, certain providers also use these linguistic varieties from outside the Union in order to avoid European Union data protection rules. The rules are much more relaxed outside the Union, and the data are obtained more through data mining.</p> | <p>nivelul de a face diferența. De exemplu, există diferențe între franceza din Franța și franceza din Quebec sau Senegal, între portugheza din Portugalia și portugheza din Brazilia sau între spaniola din Spania și spaniola din Mexic sau Argentina. Din ce am înțeles, anumiți producători folosesc aceste varietăți lingvistice din afara Uniunii și pentru a evita regulile Uniunii Europene privind protecția datelor. Regulile sunt mult mai relaxate în afara Uniunii, iar datele sunt luate mai mult prin data mining.</p> |
| <p><b>Interviewer:</b> Since you used commercial resources, how did you manage to adapt them from a terminological point of view?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | <p><b>Interviewator:</b> Din moment ce ați folosit resurse comerciale, cum ați reușit să le adaptați din punct de vedere terminologic?</p>                                                                                                                                                                                                                                                                                                                                                                                            |
| <p><b>Interviewee:</b> Through customisation. In Parliament, many acronyms, proper names and specific terms are used that do not appear in other contexts. For example, the names of DGs. In English, ENVI means envy, but for us it is the ENVI committee; PETI may sound like petty in English, but for us it is the PETI committee. All these terms need to be customised and trained into the AI.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                               | <p><b>Interviewat:</b> Prin customizare. În Parlament se folosesc foarte multe acronime, nume proprii și termeni specifici care nu apar în alte contexte. De exemplu, nume de DG-uri. ENVI în engleză înseamnă invidie, dar la noi este comisia ENVI; PETI poate suna ca petty în engleză (meschin), dar la noi este comisia PETI. Toți acești termeni trebuie customizați și antrenați în AI.</p>                                                                                                                                    |
| <p><b>Interviewer:</b> I see, so you took the commercial resources and customised them within your Unit.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | <p><b>Interviewator:</b> Am înțeles, deci ați luat resursele comerciale și le-ați customizat în interiorul Unității dumneavoastră.</p>                                                                                                                                                                                                                                                                                                                                                                                                |
| <p><b>Interviewee:</b> Yes. We are in the process of customising and training AI-based language models.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | <p><b>Interviewat:</b> Da. Suntem în proces de customizare și antrenare a modelelor lingvistice bazate pe AI (AI language models).</p>                                                                                                                                                                                                                                                                                                                                                                                                |
| <p><b>Interviewer:</b> How did you do that? With the help of translators, for all European languages?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | <p><b>Interviewator:</b> Cum ați făcut asta? Cu ajutorul traducătorilor, pentru toate limbile europene?</p>                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <p><b>Interviewee:</b> In order to save the human resources of DG TRAD Directorate, which are already overloaded with work, we relied mainly on trainees and staff from Directorate D. However, for higher-level or very delicate tasks, we asked for the support of the Parliament's professional linguists.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | <p><b>Interviewat:</b> Pentru a economisi resursele umane deja supraîncărcate de muncă ale Directoratului DG TRAD, am folosit în principal stagiaři și personal din cadrul Directoratului D. Dar, pentru sarcini de nivel mai înalt sau foarte delicate, am cerut sprijinul lingviștilor profesioniști din cadrul Parlamentului.</p>                                                                                                                                                                                                  |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Interviewer:</b> From your point of view, what are the main advantages of using Automatic Speech Recognition in an institutional context?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | <p><b>Interviewator:</b> Din punctul dumneavoastră de vedere, care sunt principalele avantaje ale utilizării Automatic Speech Recognition (recunoașterii automate a vorbirii) într-un context instituțional?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| <p><b>Interviewee:</b> The first is accessibility. There are 52 million citizens of the European Union who have disabilities that affect hearing and participation in the democratic debate process in the European Parliament. We mainly refer to persons with hearing impairments, but there are also other disabilities, for example ADHD and other disorders that make following the democratic process very difficult. This is the first advantage and our main goal, as it concerns an important proportion of the European Union's population.</p> <p>Secondly, there are practical advantages for secretariats, staff and Members of the European Parliament. We have had interviews with Members who said that if they are networking at a large meeting and do not follow what has been discussed, they can afterwards look at the transcript. Others, who are members of two or three committees, can sit physically in one committee and follow remotely what is being discussed in another. Or, on the way to the airport, they can follow on their phone what is being debated without turning on the sound. This is particularly relevant for the European Parliament because, unlike national parliaments, many events take place simultaneously: committees, subcommittees, temporary committees, full committees, group meetings. There may be four, five, six or seven events on the same day.</p> | <p><b>Intervievat:</b> Primul este accesibilitatea. Există 52 de milioane de cetățeni ai Uniunii Europene care au dizabilități care afectează auzul și participarea la procesul de dezbatere democratică din Parlamentul European. Ne referim în principal la persoanele cu deficiență de auz, dar există și alte dizabilități, de exemplu ADHD și alte tulburări care fac urmărirea procesului democratic foarte dificilă. Acesta este primul avantaj și principalul nostru scop, fiind un procent important din populația Uniunii Europene.</p> <p>În al doilea rând, există avantaje practice, pentru secretariate, pentru personal și pentru europarlamentari. Am avut interviuri cu europarlamentari în care ziceau că, dacă fac networking într-o reuniune mare și nu urmăresc ce s-a discutat, după aceea se pot uita pe transcript. Alții, care sunt membri în două sau trei comisii, pot sta într-o comisie fizic și urmări de la distanță ce se discută în cealaltă. Sau, în drum spre aeroport, pot urmări pe telefon ce se dezbate fără să pornească sunetul. Acest lucru este relevant mai ales pentru Parlamentul European deoarece, spre deosebire de parlamentele naționale, au loc multe evenimente simultan: comisii, subcomisii, temporary committees, full committees, group meetings. Pot fi patru, cinci, șase sau șapte evenimente în aceeași zi.</p> |
| <p><b>Interviewer:</b> So, even if we wanted to introduce interpreters in every session or presentation, it simply would not be possible because of the large number of events?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | <p><b>Interviewator:</b> Deci, chiar dacă am vrea să introducem interpreți în fiecare sesiune sau prezentare, nu ar fi posibil pur și simplu din cauza numărului mare de evenimente?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <p><b>Interviewee:</b> Yes. And for staff there is the practical advantage of having access to a transcript when they miss something while taking notes, whether because of noise, microphone interruption, loss of attention, etc. For example, I was in a debate and a German Member began speaking about Romania and about car insurance companies that had gone bankrupt. I did not listen from the beginning because I did not know he was going to talk</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | <p><b>Intervievat:</b> Da. Iar pentru personal există avantajul practic de a avea acces la un transcript atunci când ratează ceva în timp ce iau notițe, fie din cauza zgomotului, a întreruperii microfonului, a pierderii atenției etc. De exemplu, eram într-o dezbatere și un deputat german a început să vorbească despre România și despre firmele de asigurări auto care au dat faliment. Nu l-am ascultat de la început, căci nu știam că voia să vorbească despre România. Nu mi-am dat seama imediat unde voia să ajungă. Până mi-</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| about Romania. I did not immediately realise where he was going with it. By the time I put on my headphones and switched to English, because I do not speak German, he was already talking about something else. With the transcription, I was able to scroll back and see that he was referring to the bankruptcy of insurance companies in Romania.                                                                                                                                                                                                                                                                                                                              | am pus căștile și am trecut pe engleză (căci nu vorbesc germană), el deja vorbea despre altceva. Cu transcrierea am putut să dau scrolling back și să văd că se referea la falimentul firmelor de asigurări din România.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| <b>Interviewer:</b> What challenges or limitations did you encounter in the process of developing or implementing Automatic Speech Recognition in the European Parliament?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | <b>Interviewator:</b> Ce provocări sau limitări ai întâmpinat în procesul de dezvoltare sau implementare a Automatic Speech Recognition în Parlamentul European?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <b>Interviewee:</b> Technical, budgetary, institutional and human – all of them. From a technical point of view, it is about the technical capacity itself, error management, ensuring sound quality, finding and customising a language model, and building a sufficiently large database, the problem being with low-resource languages.                                                                                                                                                                                                                                                                                                                                         | <b>Intervievat:</b> Tehnice, bugetare, instituționale și umane – toate. Din punct de vedere tehnic, este vorba despre capacitatea tehnică propriu-zisă, gestionarea erorilor, asigurarea sunetului, găsirea și customizarea unui model lingvistic și construirea unei baze de date suficient de mari (problema fiind la low-resource languages).                                                                                                                                                                                                                                                                                                                                                                       |
| <b>Interviewer:</b> Would you not place Romanian in the category of low-resource languages for language data collection?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | <b>Interviewator:</b> Nu ai încadra româna în categoria limbilor cu resurse reduse pentru language data collection?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Interviewee:</b> No. It is surprisingly good. It depends very much on accent. We ran tests with different Members of the European Parliament, using their speeches, and it turned out very, very well – surprisingly well. I was not expecting that.                                                                                                                                                                                                                                                                                                                                                                                                                            | <b>Intervievat:</b> Nu. Este surprinzător de bună. Depinde foarte mult de accent. Am făcut teste cu diferiți europarlamentari, luând discursurile lor, și a ieșit foarte, foarte bine – surprinzător de bine. Nu mă așteptam la asta.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <b>Interviewer:</b> So, regarding Romanian, if someone comes and speaks with a strong Moldavian accent, for example, are there difficulties in speech recognition?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | <b>Interviewator:</b> Deci, în privința limbii române, dacă vine cineva și vorbește cu un puternic accent moldovenesc, de exemplu, există dificultăți în recunoașterea vorbirii?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <b>Interviewee:</b> In plenary, when Romanian Members of the European Parliament speak, this is also a cultural issue. In Romanian culture there is no incentive for MEPs to speak at the rostrum in a local dialect. On the contrary, the Romanian cultural mindset is to speak, in official situations and especially at the Parliament's rostrum, in Romanian that is as standard and as correct as possible. In general, in plenary and in committees, Romanian is fairly standardised, clear and close to the standard language.<br>It is very different, for example, in the Swiss Parliament, where people may speak much more strongly in their local variety. In the case | <b>Intervievat:</b> În plen, când vorbesc europarlamentarii români, aceasta este și o chestiune culturală. În cultura română nu există un stimulent pentru europarlamentari să vorbească la tribună în dialect local. Din contră, mentalitatea culturală română este de a vorbi, în situații oficiale și mai ales la tribuna Parlamentului, într-o română cât mai standard și cât mai corectă. În general, în plen și în comisii, limba română este destul de standardizată, clară și apropiată de limba standard.<br>Este foarte diferit, de exemplu, în Parlamentul Elvețian, unde oamenii pot vorbi mult mai pronunțat în varietatea lor locală. În cazul românei, această problemă este mult mai puțin pronunțată. |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| of Romanian, this problem is far less pronounced.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>Interviewer:</b> Are there any other relevant details about Romanian – limitations or difficulties that you encountered in customising the language model?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | <b>Interviewator:</b> Mai există alte detalii relevante despre limba română – limitări sau dificultăți pe care le-ați întâlnit în customizarea modelului lingvistic?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <p><b>Interviewee:</b> We have not had the necessary resources to devote that much attention to Romanian. Because of the limited resources we have, and because we are a small unit, we have been very busy getting the system up and running. Personally, most of my time and energy goes into the institutional and budgetary side, not the technical one. I struggle every day with institutional and budgetary issues. Therefore, we have not yet reached the point where we can customise Romanian in such detail and systematically, thoroughly check and record the errors.</p> <p>We carried out a general check of the error rate to see whether it falls within the standard, and Romanian was among the languages with the lowest error rates. Both the data used in our module, where we used standard Romanian from speeches in Parliament, and the Microsoft Azure data were actually quite good.</p> | <p><b>Interviewat:</b> Nu am avut resursele necesare ca să acordăm atât de multă atenție limbii române. Din cauza resurselor limitate pe care le avem și pentru că suntem o unitate mică, am fost foarte ocupați să pornim sistemul. Personal, cea mai mare parte din timp și energie merge în partea instituțională și bugetară, nu în cea tehnică. Mă lupt în fiecare zi cu chestiunile instituționale și bugetare. Prin urmare, nu am ajuns încă la punctul în care să customizăm româna atât de detaliat și să verificăm și să notăm sistematic, amănunțit, erorile.</p> <p>Am făcut o verificare generală a ratei de erori ca să vedem dacă se încadrează în standard, iar româna era printre limbile cu cele mai mici rate de erori. Atât datele folosite în modulul nostru, unde am folosit româna standard din discursurile din Parlament, cât și datele Microsoft Azure au fost chiar bune.</p> |
| <b>Interviewer:</b> Regarding human resources, you said that you collaborated with trainees?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | <b>Interviewator:</b> Legat de resursele umane, ați spus că ați colaborat cu stagiarii?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <b>Interviewee:</b> Yes, from across DG TRAD. We had 53 trainees from all 24 languages. Their contribution varied enormously: some completed 100% of the tasks in a few weeks, while others did 0%. For Romanian, the contribution was rather limited, because there should have been two trainees per language, but the Romanian translation unit had only one.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | <b>Interviewat:</b> Da, din tot DG TRAD. Am avut 53 de stagiaari din toate cele 24 de limbi. Contribuția lor a fost extraordinar de diferită: unii au făcut 100% din task-uri în câteva săptămâni, iar alții au făcut 0%. La română, contribuția a fost destul de redusă, pentru că ar fi trebuit să fie doi stagiaari pe fiecare limbă, însă unitatea română de traducere a avut doar unul.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| <b>Interviewer:</b> I see. Were any experiments carried out to compare a fully automated ASR workflow for live transcription with one that uses ASR together with a person who actively corrects the transcription, like a respeaker?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | <b>Interviewator:</b> Am înțeles. Au fost realizate experimente pentru a compara un flux de lucru cu ASR complet automatizat pentru transcrierea în direct cu unul care folosește ASR împreună cu o persoană care corectează activ transcrierea, ca un respeaker?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <b>Interviewee:</b> We do not have the resources. People outside Parliament find it hard to understand that, even though there are many people here, you still do not have the necessary resources. The Speech-to-Text Unit has three people on the technical side, one administrative                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | <b>Interviewat:</b> Nu avem resursele. Oamenii din afara Parlamentului înțeleg greu că, deși sunt mulți oameni aici, tot nu ai resursele necesare. Unitatea Speech-to-Text are trei persoane pe partea tehnică, o persoană administrativă și o persoană pe comunicare, plus pe mine ca șef de unitate. Suntem                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>person and one person for communication, plus me as head of unit. We are all overloaded. And if we do something, we have to do it for 24 languages.</p> <p>So, as I was saying, Romanian falls into a middle category in terms of the attention given to it. We have three broad categories. First, English, because many of the events we covered (Anti-harassment Conference, Coming Out Conference, Multilingualism Day, Clear Language Day) were 90% in English. So the customisation efforts for English are category 1A.</p> <p>Then we also try to work on French and German for events; let us say those are category 1B. At the other end are Maltese and Irish, where commercial providers have almost nothing and we are trying to save these languages so that we can use them. The other languages, including Italian, Spanish, Polish and Romanian, fall into the middle category.</p> <p>We simply have not had the time or resources to work on them very much, because we have been very busy launching the application and dealing with institutional, interinstitutional, technical and data protection issues. My colleagues have spent, I think, 200–300 working hours on data protection.</p> | <p>cu toții suprasolicitați. Iar dacă facem ceva, trebuie să facem pentru 24 de limbi.</p> <p>Deci, cum ziceam: româna se află într-o categorie de mijloc în ceea ce privește atenția acordată. Avem trei categorii mari. Mai întâi engleza, pentru că multe dintre evenimentele pe care le-am acoperit (Anti-harassment Conference, Coming Out Conference, Multilingualism Day, Clear Language Day) au fost în proporție de 90% în engleză. Deci eforturile de customizare pentru limba engleză sunt de categoria 1A.</p> <p>Apoi încercăm să facem și franceză și germană pentru evenimente; să spunem că acestea sunt categoria 1B. La celălalt capăt sunt malteza și irlandeza, unde producătorii comerciali nu au aproape nimic și noi încercăm să salvăm aceste limbi ca să le putem folosi. Celelalte limbi, inclusiv italiana, spaniola, polona și româna, intră în categoria de mijloc.</p> <p>Pur și simplu nu am avut timpul sau resursele să ne ocupăm foarte mult de ele, pentru că am fost foarte ocupați cu lansarea aplicației și cu chestiile instituționale, interinstituționale, tehnice și de protecție a datelor. Colegii mei au consumat, cred, 200-300 de ore de lucru pe data protection.</p> |
| <p><b>Interviewer:</b> And now the application continues to function by granting access to participants in certain parliamentary events on that day; it is not a public application that can be used at any time, correct?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | <p><b>Interviewator:</b> Iar acum aplicația continuă să funcționeze prin acordarea accesului participanților la anumite evenimente parlamentare în ziua respectivă; nu este o aplicație publică ce poate fi folosită oricând, corect?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| <p><b>Interviewee:</b> It is designed to be live. It is intended for participants following the debates during the event. Afterwards, it can no longer be accessed.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | <p><b>Interviewat:</b> Este gândită să fie live. Este destinată participanților care urmăresc dezbaterile în timpul evenimentului. După aceea, nu se mai poate accesa.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| <p><b>Interviewer:</b> Final question: what would you change or improve in this field of Automatic Speech Recognition in Parliament if there were no technical or financial constraints at all?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | <p><b>Interviewator:</b> Întrebare finală: ce ați schimba sau îmbunătăți în acest domeniu al Automatic Speech Recognition din Parlament dacă nu ar exista deloc constrângeri tehnice sau financiare?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <p><b>Interviewee:</b> There are two major problems. The first concerns resources for testing, investment and customisation. For example, if we want to test and verify quality, data protection issues immediately arise and data protection authorisation is needed. For some non-public events, we did not even have authorisation to save the transcript afterwards</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | <p><b>Interviewat:</b> Sunt două probleme mari. Prima este legată de resurse pentru testare, investiții și customizare. De exemplu, dacă vrem să testăm și să verificăm calitatea, apar imediat probleme de protecție a datelor și este nevoie de data protection authorisation. Pentru unele evenimente nepublice nici măcar nu am avut autorizația de a salva transcriptul după aceea și de a face un quality check</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>and carry out a proper quality check. We were only able to follow and verify it live within the Unit, each colleague for their own working languages.</p> <p>Now, because we covered a public ECON meeting broadcast live, which does not create data protection problems, we want to send the material, the transcripts, to all 24 linguistic units so that they can check word error rate, latency and overall performance. But this consumes a great deal of human resources and time. It is difficult both institutionally and logistically to organise these checks with the different linguistic units.</p> <p>The institutional complexity of a place like the European Parliament is very great. Added to this is a very strong aversion to risk. Members of the European Parliament are very risk-averse regarding the impact that an instantaneous translation containing errors could have on their speeches.</p> <p>That is why learning by doing is essential: first you do it, then you correct it, because there are things that you simply cannot anticipate in advance.</p> | <p>adevărat. Am putut urmări și verifica doar live în cadrul Unității, fiecare coleg pentru propriile limbi de lucru.</p> <p>Acum, pentru că am acoperit o reuniune ECON publică transmisă live, care nu creează probleme din punctul de vedere al protecției datelor, vrem să trimitem materialul (transcripturile) tuturor celor 24 de unități lingvistice ca să verifice word error rate, latența și performanța generală. Dar acest lucru consumă foarte multe resurse umane și timp. Este dificil din punct de vedere instituțional, cât și logistic, să organizăm aceste verificări cu diferitele unități lingvistice.</p> <p>Complexitatea instituțională a unui loc precum Parlamentul European este foarte mare. Se adaugă și o aversiune foarte puternică față de risc. Europarlamentarii sunt foarte reticenți la risc în ceea ce privește impactul pe care o traducere instantanee cu erori l-ar putea avea asupra discursurilor lor. De aceea, learning by doing este esențial: mai întâi faci, apoi corectezi, pentru că există lucruri pe care pur și simplu nu le poți anticipa dinainte.</p> |
| <p><b>Interviewer:</b> So the final aim would be the creation of an internal language model used for this application?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | <p><b>Interviewator:</b> Deci scopul final ar fi crearea unui model lingvistic intern folosit pentru această aplicație?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <p><b>Interviewee:</b> Yes, but that has also become complicated, because an internal model would be more expensive and has led to institutional debates with the IT side. Creating such an internal ASR language model would cost a great deal. So overall, besides the problem of resources for testing and customisation, there is also the institutional side – data protection, misunderstandings between DGs, competences, and the specific risk of false friends or misleading formulations in important political statements.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | <p><b>Interviewat:</b> Da, dar și acest lucru a devenit complicat, pentru că un model intern ar fi mai scump și a dus la dezbateri instituționale cu partea de IT. Crearea unui astfel de model lingvistic ASR intern ar costa foarte mult. Așadar, per total, pe lângă problema resurselor pentru testare și customizare, există și partea instituțională – protecția datelor, neînțelegerile dintre DG-uri, competențele și riscul specific al false friends sau al formulărilor înșelătoare în declarații politice importante.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <p><b>Interviewer:</b> So, in any case, any transcription produced in Parliament would need post-editing before it could be published somewhere or taken up by the press?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | <p><b>Interviewator:</b> Deci, în orice caz, orice transcriere din cadrul Parlamentului ar avea nevoie de post-editing înainte să poată fi publicată undeva sau preluată de presă?</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <p><b>Interviewee:</b> That is the problem. If you want accessibility for persons with disabilities, you have to be live. We are caught between two</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | <p><b>Interviewat:</b> Asta este problema. Dacă vrei accesibilitate pentru persoanele cu dizabilități, trebuie să fii live. Noi suntem prinși între două</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>contradictory things. On the one hand, we want quality and accuracy; on the other hand, persons with disabilities want the text to be identified, transcribed and translated within a maximum of three seconds. That is their standard.</p> <p>It may happen that an MEP says that the bombing of Gaza is very bad and that millions of people have died, but if the application misses the words very bad, the text may appear as if the MEP supports the bombing. The same applies in the case of other sensitive topics: climate change, Trump, or other political subjects. Take today's conversation as an example: I switched many times from Romanian into English and back again. The application can get confused. If I say library in English while we are speaking Romanian (Let's meet at the library), the ASR application may render librărie instead of bibliotecă. And that is also risky for MEPs, because they too do switch between languages.</p> <p>In an academic research environment, people mainly look at quality and work to make an application good. Here, at the same time, we have to make sure we do not upset the interpreters, that Members do not feel exposed if something politically sensitive is rendered incorrectly, and that we manage the budgetary implications and the division between digital interpretation and digital translation. All this creates a much more complex environment, so that the focus on quality and implementation, as well as the resources we can devote to this part, are reduced by all the other constraints. We are more Catholic than the Pope when it comes to data protection.</p> | <p>lucruri contradictorii. Pe de o parte vrem calitate și acuratețe; pe de altă parte, persoanele cu dizabilități vor ca textul să fie identificat, transcris și tradus în maximum trei secunde. Acesta este standardul lor.</p> <p>Se poate întâmpla ca un europarlamentar să spună că bombardamentul din Gaza este foarte rău și că au murit milioane de oameni, dar dacă aplicația ratează cuvintele foarte rău, textul poate apărea ca și cum europarlamentarul ar susține bombardamentul. Același lucru se aplică și în cazul altor subiecte sensibile: schimbările climatice, Trump sau alte subiecte politice. Luați drept exemplu conversația noastră de astăzi: eu am trecut de multe ori din română în engleză și înapoi. Aplicația se poate zăpăci. Dacă spun library în engleză în timp ce vorbim în română (Hai să ne vedem la library), aplicația ASR poate reda librărie în loc de bibliotecă. Și asta este riscant și pentru europarlamentari, pentru că și ei fac switching between languages.</p> <p>Într-un mediu universitar de cercetare, oamenii se uită în principal la calitate și lucrează pentru ca o aplicație să fie bună. Noi, aici, trebuie, în același timp, să avem grijă să nu supărăm interpreții, ca parlamentarii să nu se simtă expuși dacă ceva sensibil politic este redat greșit, și să gestionăm implicațiile bugetare și împărțirea dintre digital interpretation și digital translation. Toate acestea creează un mediu mult mai complex, astfel încât focusul pe calitate și implementare, precum și resursele pe care le putem acorda acestei părți, sunt reduse de toate celelalte constrângeri. Suntem mai catolici decât Papa în privința data protection.</p> |
| <p><b>Interviewer:</b> I understand. I think it is also because, for plenary sessions, Verba is already used in an active way – someone has to sit there and correct it, so that something the Member did not intend to say does not come out.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | <p><b>Interviewator:</b> Am înțeles. Cred că este și din cauza faptului că, pentru sesiunile plenare, Verba este deja folosită într-un mod activ – trebuie să stea cineva acolo și să corecteze, ca să nu reiasă ceva ce parlamentarul nu ar fi vrut să spună.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| <p><b>Interviewee:</b> Exactly. And that is why, for example, at plenary level, DG Presidents do not want to use our application because of this risk. But Verba is not live and nor is it automatically translated into all 24 languages, so it is not accessible for persons with disabilities.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | <p><b>Intervievat:</b> Exact. Și de aceea, de exemplu, la nivel de plen DG Presidents nu vor să folosească aplicația noastră din cauza acestui risc. Dar Verba nu este live și nici nu este tradusă automat în toate cele 24 de limbi, deci nu este accesibilă pentru persoanele cu dizabilități.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

**Interviewer:** Thank you very much, and I wish you continued success with this application!

**Interviewator:** Vă mulțumesc și vă urez în continuare succes cu această aplicație!



## Appendix B - Interview on Verba (EP RO Unit, DG TRAD)

### Interview questions

| No. | English                                                                                                                                          | Romanian                                                                                                                         |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
|     | <b>General Questions</b>                                                                                                                         | <b>Întrebări Generale</b>                                                                                                        |
| 1   | Could you explain what Verba is and how you use it in your workflow as a Romanian proofreader at the European Parliament?                        | Ai putea explica ce este Verba, cum și când o folosești în calitate de proofreader de limba română la Parlamentul European?      |
| 2   | Do you feel that using Verba improves the speed or efficiency of your transcription work?                                                        | Consideri că, mulțumită aplicației Verba, procesul de transcriere este mai rapid și eficient?                                    |
| 3   | How accurate is Verba in transcribing Romanian speeches, particularly in the context of plenary sessions?                                        | Cât de precisă este Verba în transcrierea discursurilor în limba română, în contextul ședințelor plenare?                        |
| 4   | How much correction do you typically need to perform on Romanian transcriptions automatically generated by Verba? What are some common mistakes? | Cât de mult trebuie să corectăm, de obicei, transcrierile în limba română generate automat de Verba? Care sunt greșelile tipice? |

### Transcript and recording

#### Recording file (RO):

<https://drive.google.com/file/d/1ry3EFJJDKXnoowmDPjGI3sjWHXCWKaHH/view?usp=sharing>

| English Translation                                                                                                                                                                                                                     | Romanian Original                                                                                                                                                                                                                         |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Interviewer:</b> So, what is Verba?                                                                                                                                                                                                  | <b>Intervievator:</b> Deci, ce este Verba?                                                                                                                                                                                                |
| <b>Interviewee:</b> It is an application through which Members' speeches are transcribed.                                                                                                                                               | <b>Intervievat:</b> Este o aplicație prin care se transcriu discursurile deputaților.                                                                                                                                                     |
| <b>Interviewer:</b> Can there also be special guests?                                                                                                                                                                                   | <b>Intervievator:</b> Pot fi și invitați speciali?                                                                                                                                                                                        |
| <b>Interviewee:</b> Yes, mainly for plenary sittings, but also in committees. For example, in the case of the Commissioners, we had the Commissioners' hearings. So there are also other situations. How do we use it? As proofreaders, | <b>Intervievat:</b> Da, deci, în principal pentru ședințele plenare, dar și în comisii. De exemplu, în cazul comisarilor, am avut audierile comisarilor. Deci sunt și alte situații. Cum o folosim noi? Noi, în calitate de proofreaderi, |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| we have to edit the speeches. We mainly use Verba during the sessions.                                                                                                                                                                                                                                                                                                                                                                                                                               | trebuie să edităm discursurile. În principal, noi folosim Verba în timpul sesiunilor.                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Interviewer:</b> So it is used not only for plenary sessions, but also in other cases, though only for high-level meetings?                                                                                                                                                                                                                                                                                                                                                                       | <b>Interviewator:</b> Se folosește nu numai în cazul sesiunilor plenare, ci și în alte cazuri, însă numai pentru reuniunile la nivel înalt?                                                                                                                                                                                                                                                                                                               |
| <b>Interviewee:</b> Exactly. We transcribe them in order to publish them officially. That is how they reach us.                                                                                                                                                                                                                                                                                                                                                                                      | <b>Intervievat:</b> Exact. Noi le transcriem ca să le publicăm oficial. Așa ajung la noi.                                                                                                                                                                                                                                                                                                                                                                 |
| <b>Interviewer:</b> Do you consider that, thanks to the Verba application, the process is faster and more efficient?                                                                                                                                                                                                                                                                                                                                                                                 | <b>Interviewator:</b> Consideri că, mulțumită aplicației Verba, procesul este mai rapid și mai eficient?                                                                                                                                                                                                                                                                                                                                                  |
| <b>Interviewee:</b> Clearly. And compared with CRE, which we used before, it is much better. Not only does it recognise the spoken language better, and you can also see a bit of the hand of AI in it. It seems to learn.                                                                                                                                                                                                                                                                           | <b>Intervievat:</b> Clar. Dar și față de CRE, pe care o foloseam înainte, e mult mai bună. Nu numai că recunoaște mai bine limba vorbită, eu recunosc aici un pic mâna AI-ului. Adică pare că învață.                                                                                                                                                                                                                                                     |
| <b>Interviewer:</b> I see, so it has improved over time.                                                                                                                                                                                                                                                                                                                                                                                                                                             | <b>Interviewator:</b> Am înțeles, deci s-a îmbunătățit în timp.                                                                                                                                                                                                                                                                                                                                                                                           |
| <b>Interviewee:</b> Exactly. And I have noticed that even in the case of speeches that are less well pronounced or formulated, it suggests solutions. It even removes some of the extra ‘ş’ sounds. Exactly as we do when correcting. Probably, as we work on the speeches, the application learns. It even knows how to place punctuation quite well, although it also makes mistakes. Quite simply, there are sentences that end where they should not end, depending on the speaker’s intonation. | <b>Intervievat:</b> Exact. Și am observat că, chiar și în cazul discursurilor mai puțin bine pronunțate sau formulate, îți propune soluții. Mai taie din ș-uri. Exact cum corectăm noi. Probabil că, pe măsură ce noi facem discursurile, aplicația învață. Chiar cunoaște, pune semnele de punctuație destul de bine, mai dă și rateuri. Pur și simplu sunt fraze care se termină unde nu ar trebui să se termine, în funcție de intonația vorbitorului. |
| <b>Interviewer:</b> Do you think that is also because the speaker makes a long pause while speaking?                                                                                                                                                                                                                                                                                                                                                                                                 | <b>Interviewator:</b> Consideri că este și din cauza faptului că vorbitorul face o pauză mare în vorbire?                                                                                                                                                                                                                                                                                                                                                 |
| <b>Interviewee:</b> Yes, but anyway Verba is much, much improved compared with what it was before. And yes, it clearly helps us. It happened to me once, although I no longer remember why, that the application did not work and I did not have the transcript of that speech already generated, so I had to write it by hand. I found it awful. And I imagined what it would have meant if we had to sit and actually write everything.                                                            | <b>Intervievat:</b> Da, dar oricum Verba e mult, mult îmbunătățită față de ce era înainte. Și da, e clar că ne ajută. Mi s-a întâmplat odată, dar nu mai știu de ce, cred că n-a funcționat aplicația și nu aveam transcriptul discursului respectiv deja generat și a trebuit să-l scriu de mână. Mi s-a părut groaznic. Și mi-am imaginat ce ar fi însemnat dacă am sta să scriem efectiv.                                                              |
| <b>Interviewer:</b> How accurate is it in the context of plenary sittings?                                                                                                                                                                                                                                                                                                                                                                                                                           | <b>Interviewator:</b> Cât de precisă este, în contextul ședințelor plenare?                                                                                                                                                                                                                                                                                                                                                                               |
| <b>Interviewee:</b> As I said, it is fairly accurate.                                                                                                                                                                                                                                                                                                                                                                                                                                                | <b>Intervievat:</b> Cum ziceam, e destul de precisă.                                                                                                                                                                                                                                                                                                                                                                                                      |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Interviewer:</b> What I meant by that is this: plenary sittings have a certain language and a certain terminology; certain committees and certain important cities are mentioned.                                                                                                                                                                                                                                                                                                                                                  | <b>Interviewator:</b> Ce am vrut să întreb prin asta este că şedinţele plenare au un anumit limbaj, o anumită terminologie; se specifică anumite comisii şi anumite oraşe importante.                                                                                                                                                                                                                                                                                                     |
| <b>Interviewee:</b> Exactly, there are all sorts of names of EU programmes. It does not really recognise those. It does recognise some of the more recurrent ones, such as Green Deal. But there are some that it simply does not know. Probably, again, as they have been used, it has recorded them. They have clearly introduced from the start the names of committees, Members, Commissioners and all that — the people from the PV-CRE Unit (Procès verbaux et comptes rendus de la plénière) — but there are also other names. | <b>Interviewat:</b> Exact, sunt tot felul de denumiri de programe UE. Pe acestea nu prea le recunoaşte. Pe unele mai recurente, da, gen Green Deal. Dar sunt unele pe care nu le ştie, efectiv. Probabil că, tot aşa, pe măsură ce au fost folosite, le-a înregistrat. Deci au introdus ei — cei de la PV-CRE Unit (Procès verbaux et comptes rendus de la plénière) — clar numele de comisii, deputaţii, comisarii; toate astea le-au introdus din start, dar mai sunt şi alte denumiri. |
| <b>Interviewer:</b> Newly appeared ones?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | <b>Interviewator:</b> Cele nou-apărute?                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <b>Interviewee:</b> Yes, ones that it does not always recognise, that is clear. But otherwise, yes, it is fairly accurate and it seems to me that it also corrects from a grammatical point of view. So very often I notice that it has corrected something.                                                                                                                                                                                                                                                                          | <b>Interviewat:</b> Da, pe cele pe care nu totdeauna le recunoaşte, asta e clar. Dar, în rest, da, e destul de precisă şi mi se pare că corectează şi din punct de vedere gramatical. Deci chiar de multe ori observ că a corectat.                                                                                                                                                                                                                                                       |
| <b>Interviewer:</b> Meaning that what the speaker said was wrong and the application corrected it?                                                                                                                                                                                                                                                                                                                                                                                                                                    | <b>Interviewator:</b> Adică ce a spus vorbitorul era greşit, iar aplicaţia a corectat?                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Interviewee:</b> Exactly, it corrected the grammatical agreement.                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | <b>Interviewat:</b> Exact, a făcut acordul gramatical.                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Interviewer:</b> What are the typical mistakes for Romanian?                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | <b>Interviewator:</b> Care sunt greşelile tipice pentru limba română?                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Interviewee:</b> I do not know if there are typical mistakes. It makes mistakes because of the speakers, let us say. Clearly, if the speaker repeats themselves, the automatic transcription also repeats it. Although sometimes I have seen that it cuts things out; where there is clearly a repetition, it removes it. If an MEP stutters, it has cut that part out, clearly.                                                                                                                                                   | <b>Interviewat:</b> Nu ştiu dacă sunt greşeli tipice. Face greşeli tot din cauza vorbitorilor, să zicem. Clar, dacă vorbitorul se repetă, se repetă şi în transcrierea automată. Deşi uneori am văzut că mai taie; acolo unde e clar o repetiţie, taie. Dacă un deputat se bâlbâie, a tăiat acea parte, clar.                                                                                                                                                                             |
| <b>Interviewer:</b> Look, we have the phonetic word ‘sau’ in Romanian, which can mean either ‘or’ or, in the past tense, for example in ‘s-au luat multe măsuri’. In this context, does it recognise the difference between ‘sau’ written as one word and ‘s-au’ written separately?                                                                                                                                                                                                                                                  | <b>Interviewator:</b> Uite, avem cuvântul fonetic „sau” în română, care poate să însemne fie conjuncţia „sau”, fie forma verbală din trecut, de exemplu în „s-au luat multe măsuri”. Recunoaşte în acest context diferenţa între „sau” scris legat şi „s-au” scris despărţit?                                                                                                                                                                                                             |
| <b>Interviewee:</b> I do not think I have ever seen a mistake of that kind.                                                                                                                                                                                                                                                                                                                                                                                                                                                           | <b>Interviewat:</b> Nu cred că am văzut niciodată o greşeală de genul ăsta.                                                                                                                                                                                                                                                                                                                                                                                                               |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Interviewer:</b> Or a similar case, a word which in Romanian is pronounced the same way but spelled differently?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | <b>Interviewator:</b> Sau un caz asemănător, un cuvânt care în română se pronunță la fel, dar se scrie diferit?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <b>Interviewee:</b> No, I have not seen that.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | <b>Interviewat:</b> Nu, n-am văzut.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Interviewer:</b> So it also recognises the context.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | <b>Interviewator:</b> Deci recunoaște și contextul.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <p><b>Interviewee:</b> There are still word confusions, but again only because of pronunciation. It understands something else.</p> <p>So yes, we do still correct. There are speeches where I correct very little. In general, when speakers read, have correct pronunciation and pronounce the sentences properly, there is almost no need for our intervention. Maybe a comma or something very minor, but that is all.</p> <p>For certain speakers who speak more freely and forget to finish one sentence before beginning a new one, our intervention is needed there.</p> <p>And then there are the mistakes that occur because of pronunciation. Not very many, but there are some. You really do have to listen carefully to the speech in order to notice the mistakes.</p> | <p><b>Interviewat:</b> Mai sunt confuzii de cuvinte, dar tot din cauza pronunției, efectiv. Înțelege altceva.</p> <p>Deci da, mai corectăm. Sunt discursuri unde corectez foarte puțin. În general, vorbitorii care citesc, care au pronunție corectă și pronunță frazele corect. Deci nu e aproape deloc nevoie de intervenția noastră. Poate vreo virgulă sau ceva infim, dar atât.</p> <p>Anumiți vorbitori care vorbesc mai liber și care uită să termine o frază și încep o alta nouă, acolo e nevoie de intervenția noastră.</p> <p>Și apoi mai sunt greșelile care apar din cauza pronunției. Nu foarte multe, dar sunt. Adică chiar trebuie să asculți cu atenție discursul ca să sesizezi greșelile.</p>       |
| <b>Interviewer:</b> Ideally, is there anything that could still be improved in the Verba application?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | <b>Interviewator:</b> În mod ideal, s-ar mai putea îmbunătăți ceva la aplicația Verba?                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <p><b>Interviewee:</b> For it to recognise everything, all the words. Of course — ironically. But that is what I think: the application improves as we work in it, that is, as we correct transcripts automatically generated with ASR.</p> <p>Clearly, it should learn the new names, or they should be introduced manually. I do not know exactly how it works.</p>                                                                                                                                                                                                                                                                                                                                                                                                                 | <p><b>Interviewat:</b> Să recunoască tot, toate cuvintele. Bineînțeles — ironic. Dar eu asta cred: aplicația se îmbunătățește pe măsură ce noi lucrăm în ea, adică pe măsură ce corectăm transcrieri generate automat cu ASR.</p> <p>Clar, denumirile noi ar trebui să le învețe sau să fie introduse manual. Nu știu exact cum funcționează.</p>                                                                                                                                                                                                                                                                                                                                                                       |
| <b>Interviewer:</b> From what I know from a previous interview, the Speech-to-Text Unit customises its ASR platform. As regards events, there is also another difference between CRE and the platform used by the Speech-to-Text Unit. They use the platform only for public events or for committee meetings that are in any case broadcast live on Parliament's website. So they use it for everything that is public, apart from these plenary sessions, which are at a far too high level, and apart from other high-level meetings. However, in the case of the Verba application, does the ASR system learn by itself or do the people from PV-CRE personally add certain specific terms and names?                                                                             | <p><b>Interviewator:</b> Din câte știu dintr-un interviu anterior, Unitatea Speech-to-Text își personalizează platforma ASR. În ceea ce privește evenimentele, este și o altă diferență între CRE și platforma celor de la Speech-to-Text Unit. Ei folosesc platforma doar pentru evenimente publice sau pentru reuniuni ale comisiei care oricum se difuzează live pe site-ul Parlamentului. Deci o folosesc pentru tot ce e public, în afară de aceste sesiuni plenare, care sunt la un nivel mult prea mare, și în afară de alte reuniuni la nivel înalt.</p> <p>Însă, în cazul aplicației Verba, sistemul ASR învață singur sau cei de la PV-CRE adaugă personal anumiți termeni specifici și anumite denumiri?</p> |

|                                                                         |                                                                 |
|-------------------------------------------------------------------------|-----------------------------------------------------------------|
| <b>Interviewee:</b> Even I do not know that. Only the PV-CRE Unit does. | <b>Intervievat:</b> Asta nici eu nu știu. Doar Unitatea PV-CRE. |
| <b>Interviewer:</b> I understand, thank you for the interview.          | <b>Intervievator:</b> Am înțeles, vă mulțumesc pentru interviu. |

## **Appendix C - Presentation for the EP RO Unit, DG TRAD**

### **Presentation video**

<https://drive.google.com/file/d/1hTN7w2ulLH6z9nU8dNxgORCurVVsf7Av/view?usp=sharing>

### **Texts used**

<https://www.europarl.europa.eu/petitions/en/petition/content/0886%252F2025/html/Petition-No-0886%252F2025-by-I.-L.-P.-%2528Romanian%2529-on-regulation-of-competitions-for-animal-traction-in-the-European-Union>

### **EN VERSION**

**0886/2025**

**Subject:   Petition No 0886/2025 by I. L. P. (Romanian) on regulation of competitions for animal traction in the European Union**

### **Summary of petition**

The petitioner refers to the fact that, during a publicly organised horse logging competition in the village of Cormaia, Bistrița-Năsăud County, Romania, a heavy draft horse collapsed and died while pulling an extremely heavy load of timber in front of spectators. According to the petitioner, this event was held legally and is part of a series of similar widespread competitions across rural Romania and other regions of Eastern Europe. The petitioner considers that there is a gap in the current EU legal framework concerning the welfare of equines used in traditional or entertainment-based traction competitions. She refers to Article 13 of the Treaty on the Functioning of the European Union, pointing out that current EU legislation fails to cover practices of public animal traction competitions, leaving them at the discretion of national or local authorities, many of which lack the capacity or willingness to enforce minimum animal welfare standards. The petitioner urges the European Parliament to take action to create a unified legal framework at EU level to prevent further similar incidents. Furthermore, the petitioner calls on the European Commission to investigate the regulatory gaps regarding such competitions in Member States, to consider introducing legislation that specifically targets the welfare of equines used in non-commercial physical labour, such as logging contests, folkloric displays or festivals- to request Romania to provide clarifications regarding the incident in Cormaia, and to start a dialogue with Member States regarding the possibility of a future ban on such events.

0886/2025

**Subiect: Petiția nr. 0886/2025, adresată de I. L. P., de cetățenie română, privind reglementarea concursurilor care presupun tracțiune animală în Uniunea Europeană**

### Rezumatul petiției

Petiționara face referire la faptul că, în cadrul unui concurs public de tras bușteni organizat în satul Cormaia, județul Bistrița-Năsăud, România, un cal de tracțiune s-a prăbușit și a murit în timp ce trăgea un buștean extrem de greu, în fața spectatorilor. Potrivit petiționarei, evenimentul s-a desfășurat legal și face parte dintr-o serie de concursuri similare larg răspândite în mediul rural din România și în alte regiuni din Europa de Est. Ea consideră că există o lacună în cadrul juridic actual al UE în ceea ce privește bunăstarea ecvideelor utilizate în concursuri de tracțiune tradiționale sau de divertisment și invocă articolul 13 din Tratatul privind funcționarea Uniunii Europene, subliniind că legislația actuală a UE nu acoperă practicile concursurilor publice cu tracțiune animală, lăsând reglementarea la aprecierea autorităților naționale sau locale, iar multe dintre acestea nu au capacitatea sau disponibilitatea de a aplica standarde minime privind bunăstarea animalelor. Ea îndeamnă Parlamentul European să ia măsuri pentru crearea unui cadru juridic unitar la nivelul UE, care să prevină astfel de incidente în viitor. Mai mult, petiționara solicită Comisiei Europene să investigheze lacunele în materie de reglementare în ceea ce privește astfel de concursuri în statele membre, să aibă în vedere introducerea unei legislații care să vizeze în mod specific bunăstarea ecvideelor utilizate în activități fizice necomerciale – cum ar fi concursurile de tras bușteni, spectacolele sau festivalurile folclorice –, să solicite României clarificări cu privire la incidentul din Cormaia și să inițieze un dialog cu statele membre cu privire la posibilitatea unei viitoare interziceri a unor astfel de evenimente.

### Explanation transcript

| English Translation                                                                                                                                                                                                                                                                                                                                | Romanian Original                                                                                                                                                                                                                                                                                                                                     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Hello everyone. As you can see on the screen, today I am going to present Automatic Speech Recognition, or ASR, also known as Speech-to-Text, or, in Romanian, recunoaștere automată a vorbirii. Before I begin the presentation, I would like to make two brief remarks.                                                                          | Salutare tuturor. După cum puteți vedea pe ecran, astăzi o să vă prezint Automatic Speech Recognition, ASR, cunoscut și sub denumirea de Speech-to-Text, sau, în limba română, recunoaștere automată a vorbirii. Înainte să încep prezentarea, aș vrea să fac două scurte precizări.                                                                  |
| First of all, I am presenting a video instead of a PowerPoint. I hope it will be more interesting and clearer. Secondly, I would like to present this technology exactly as I see it: as an option to speak instead of typing, not as the idea of replacing something or someone, as people often fear in discussions about AI. It is simply a way | În primul rând, o să vă prezint un video în locul unui PowerPoint. Sper să fie mai interesant și mai clar. În al doilea rând, aș vrea să vă prezint această tehnologie exact așa cum o văd eu: ca pe o opțiune de a vorbi în loc de a tasta, nu ca pe ideea de a înlocui ceva sau pe cineva, așa cum apare adesea teama în discuțiile despre AI. Este |

|                                                                                                                                                                                                                                                                                                                                                                                                                               |                                                                                                                                                                                                                                                                                                                                                                                                                        |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| in which we could simplify certain tasks, both in our private and professional lives. But let us look at it in more detail.                                                                                                                                                                                                                                                                                                   | doar o modalitate prin care ne-am putea simplifica anumite sarcini, atât în viața privată, cât și în cea profesională. Dar haideți să vedem mai în detaliu.                                                                                                                                                                                                                                                            |
| I also included a short video recorded in the car, to show you how we can use automatic speech recognition in everyday life. For example, in the car we can dictate to Siri, or to the voice assistant already integrated into the car, the destination we want to go to: “Hey Siri, directions to the European Parliament.”                                                                                                  | Am inclus și un scurt video filmat în mașină, pentru a vă arăta cum putem folosi recunoașterea automată a vorbirii în viața de zi cu zi. De exemplu, în mașină putem dicta asistentului vocal Siri, sau asistentului vocal integrat deja în mașină, direcția în care vrem să mergem: „Hey Siri, directions to the European Parliament”.                                                                                |
| But I would like to go into a bit more detail regarding this technology, automatic speech recognition, or Speech-to-Text. It is a technology that converts spoken language into text, so that we can interact with different virtual assistants, use transcription services, as in our case with Verba, or use it as an accessibility tool.                                                                                   | Dar aș dori să intrăm puțin mai în detaliu în privința acestei tehnologii, recunoașterea automată a vorbirii, sau Speech-to-Text. Este o tehnologie care convertește limbajul vorbit în text, astfel încât să putem interacționa cu diferiți asistenți virtuali, să putem folosi servicii de transcriere, cum este în cazul nostru Verba, sau să o utilizăm ca instrument de accesibilitate.                           |
| In short, the process behind this technology is the following: the system captures sound waves, processes them in order to extract acoustic features and patterns, then interprets them with the help of acoustic and language models, and finally decodes them.                                                                                                                                                              | Pe scurt, procesul din spatele acestei tehnologii este următorul: sistemul captează undele sonore, le procesează astfel încât să poată extrage trăsături și tipare acustice, apoi le interpretează cu ajutorul unor modele acustice și lingvistice, iar în final le decodifică.                                                                                                                                        |
| The result can be written text, meaning that we can see how what we say is turned into text, or it can be used by voice assistants to trigger a specific reaction or action.                                                                                                                                                                                                                                                  | Rezultatul poate fi un text scris, adică vedem cum ceea ce spunem se transformă în text, sau poate fi folosit pentru asistenți vocali, pentru a declanșa o anumită reacție sau o anumită acțiune din partea lor.                                                                                                                                                                                                       |
| As I mentioned, Automatic Speech Recognition is also used as an accessibility tool. The Speech-to-Text Unit within the European Parliament has developed a platform that can be accessed when we attend public events of the Parliament.                                                                                                                                                                                      | După cum v-am spus, Automatic Speech Recognition este folosit și ca instrument de accesibilitate. Unitatea Speech-to-Text din cadrul Parlamentului European a dezvoltat o platformă care poate fi accesată atunci când participăm la evenimente publice ale Parlamentului.                                                                                                                                             |
| I am also going to show you a more concrete example related to this unit and this platform. As you can see, there is a sign language interpreter, but at the same time, as we can observe in the case of this participant who is wearing a hearing aid, this platform created by the Speech-to-Text Unit can also be used. This is the platform I am referring to; the languages can be changed here, on the right-hand side. | O să vă arăt și un exemplu mai concret legat de această unitate și de această platformă. După cum vedeți, există un interpret în limbaj mimico-gestual, dar, în același timp, după cum observăm la acest participant care poartă aparat auditiv, se poate folosi și această platformă creată de Unitatea Speech-to-Text. Aceasta este platforma despre care vorbesc; limbile pot fi schimbate aici, în partea dreaptă. |
| So, it is not only a tool that transcribes what the speaker says, but also a tool that translates in real time. I personally attended an event in which this tool was used, at the PETI committee, and I                                                                                                                                                                                                                      | Așadar, nu este doar un instrument care transcrie ceea ce spune vorbitorul, ci și un instrument care traduce în timp real. Eu însămi am participat la un eveniment în care a fost folosit acest                                                                                                                                                                                                                        |



|                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| found that it translated quite well and, above all, transcribed very well in real time.                                                                                                                                                                                                                                                                                   | instrument, la comisia PETI, și mi s-a părut că traduce destul de bine și, mai ales, că transcrie foarte bine în timp real.                                                                                                                                                                                                                                                        |
| The language model used in this case was adapted by the Speech-to-Text Unit so that it would match parliamentary discourse, where different committees, proper names and specific terms appear.                                                                                                                                                                           | Modelul lingvistic folosit în acest caz a fost adaptat de către Unitatea Speech-to-Text pentru a corespunde discursului parlamentar, în care apar diferite comisii, nume proprii și termeni specifici.                                                                                                                                                                             |
| Now let us move a little further and look at the simplest method of using automatic speech recognition on any computer running Windows 10 or 11, namely the Windows Voice Typing function, which can be activated by pressing the Windows key and the H key at the same time.                                                                                             | Dar haideți să mergem puțin mai departe și să vă arăt cea mai simplă metodă de a folosi recunoașterea automată a vorbirii pe orice calculator care are Windows 10 sau 11, și anume funcția Voice Typing de la Windows, care se poate activa prin apăsarea simultană a tastei Windows și a tastei H.                                                                                |
| This option comes with different voice commands, which can be dictated in Romanian or in other languages. It also has a few basic voice functions. I am going to show you how useful they can be, even though they are simple, and I am going to dictate an entire email using only this automatic speech recognition function.                                           | Această opțiune vine cu diferite comenzi vocale, care pot fi dictate în română sau în alte limbi. Are și câteva funcții vocale de bază. O să vă arăt cât de utile pot fi, chiar dacă sunt simple, și o să vă dictez un e-mail întreg doar prin intermediul acestei funcții de recunoaștere automată a vorbirii.                                                                    |
| Let us see. This is the text box that appears when we press these two keys simultaneously. “Good afternoon, comma. New paragraph, new paragraph. Sorry to bother you.”                                                                                                                                                                                                    | Să vedem. Aceasta este căsuța de text care se activează în momentul în care apăsăm aceste două taste simultan. „Bună ziua, virgulă. Paragraf nou, paragraf nou. Mă scuzați de deranj.”                                                                                                                                                                                             |
| As you can see, I used the backspace command here because it did not automatically insert the question mark, so I had to delete and dictate the question mark myself so that it would appear correctly. However, overall, I found that it worked quite well. We can also take notes in Word, and we can disable automatic punctuation so that it does not get in the way. | După cum puteți vedea, am folosit comanda backspace aici, pentru că nu mi-a introdus automat semnul întrebării, așa că a trebuit să șterg și să dictez eu semnul întrebării, pentru ca acesta să apară corect. Însă, per ansamblu, mi s-a părut că a funcționat destul de bine. De asemenea, putem lua notițe în Word și putem dezactiva punctuația automată, ca să nu ne încurce. |
| So, with this ASR function, Windows Voice Typing, we can dictate in any text field where we would normally also be able to type on a Windows 10 or 11 laptop. However, there is one problem: this function is disabled on Parliament laptops because it relies on the cloud. Therefore, for data protection reasons, it cannot be used on institutional devices.          | Așadar, prin această funcție ASR, Voice Typing de la Windows, putem dicta în orice câmp de text în care am putea și tasta pe un laptop cu Windows 10 sau 11. Există însă o problemă: această funcție este dezactivată pe laptopurile Parlamentului, deoarece se bazează pe cloud. Așadar, din motive de protecție a datelor, ea nu poate fi folosită pe dispozitivele instituției. |
| In this context, I asked Parliament’s Service Desk what alternatives exist. How can we use automatic speech recognition on Parliament laptops, and not only on our personal devices?                                                                                                                                                                                      | În acest context, am întrebat Service Desk-ul Parlamentului ce alternative există. Cum putem folosi recunoașterea automată a vorbirii pe laptopurile Parlamentului, și nu doar pe laptopurile noastre personale?                                                                                                                                                                   |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Dragon NaturallySpeaking would be one possible solution. However, at the moment, it is not available in Romanian, and access to licences is limited. Parliament can offer employees the possibility to try this speech recognition option for six months, as I was also able to do, but for an unlimited licence a medical certificate is required to confirm the need to use this application.                                                       | Dragon NaturallySpeaking ar fi o soluție, însă, momentan, nu este disponibil în limba română, iar accesul la licențe este limitat. Parlamentul le poate oferi angajaților posibilitatea de a încerca această opțiune de recunoaștere a vorbirii timp de șase luni, așa cum am avut și eu ocazia, însă pentru o licență pe termen nelimitat este nevoie de un certificat medical care să ateste necesitatea utilizării acestei aplicații. |
| But who knows? Perhaps Romanian will also be integrated in the future, considering that there are already enough ASR data and resources for Romanian. In addition, if there is feedback from people who try this application, perhaps more licences could be made available. And even if not, it is still worth testing for six months.                                                                                                               | Dar cine știe? Poate, pe viitor, va fi integrată și limba română, având în vedere că există deja destule date și resurse ASR pentru limba română. În plus, dacă va exista feedback din partea celor care ar încerca această aplicație, poate se vor putea oferi și mai multe licențe. Iar dacă nu, măcar merită testată timp de șase luni.                                                                                               |
| I am also going to show you my attempt to use Dragon NaturallySpeaking. As you can see, it has to be installed separately; it is not already integrated into another program. A small microphone icon appears, just as in the case of Microsoft Voice Typing. Different voice commands can be used and, in addition, audio files can also be transcribed. We can upload them and then, as you can see, the transcribed file is created automatically. | O să vă arăt și încercarea mea de a folosi Dragon NaturallySpeaking. După cum puteți vedea, trebuie instalat separat, nu este deja integrat într-un alt program. Apare o iconiță cu un microfon, la fel ca în cazul Microsoft Voice Typing. Se pot folosi diferite comenzi vocale și, în plus, se pot transcrie și fișiere audio. Le putem încărca, iar apoi, după cum puteți vedea, ni se creează automat fișierul transcris.           |
| However, what I would really like to show you, as translators, is how we can use this function not only to dictate an email, not only to take notes, not only in private life, but also within the translation process.                                                                                                                                                                                                                               | Însă ceea ce aș vrea să vă arăt vouă, traducătorilor, este cum putem folosi această funcție nu doar pentru a dicta un e-mail, nu doar pentru a lua notițe, nu doar în viața privată, ci și în cadrul procesului de traducere.                                                                                                                                                                                                            |
| Dragon NaturallySpeaking and Voice Typing, although Voice Typing unfortunately cannot be used on Parliament laptops, also work in the translation environment of Trados Studio, desktop version. I am going to show you an example of translation in a Trados Studio segment, unfortunately from Romanian into English, because the Romanian option is not yet available in all the necessary configurations.                                         | Dragon NaturallySpeaking și Voice Typing, deși Voice Typing, din păcate, nu poate fi folosit pe laptopurile Parlamentului, funcționează și în mediul de traducere Trados Studio, varianta desktop. O să vă arăt un exemplu de traducere într-un segment din Trados Studio, din păcate din română în engleză, pentru că opțiunea pentru limba română nu este încă disponibilă în toate configurațiile necesare.                           |
| Still, the best solution with regard to automatic speech recognition in the translation environment, at least for CAT tools and at least up to 2024, is Trados Online Editor and its integrated voice dictation function.                                                                                                                                                                                                                             | Totuși, cea mai bună soluție în ceea ce privește recunoașterea automată a vorbirii în mediul de traducere, cel puțin pentru CAT tools și cel puțin până în 2024, este Trados Online Editor și funcția sa integrată de dictare vocală.                                                                                                                                                                                                    |
| It includes many more voice commands that can be used specifically for translation work. I know that we do not use Trados Online Editor, but rather the desktop version; however, considering                                                                                                                                                                                                                                                         | Acesta include mult mai multe comenzi vocale, care pot fi folosite special pentru activitatea de traducere. Știu că noi nu folosim Trados Online Editor, ci varianta desktop, dar având în vedere                                                                                                                                                                                                                                        |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| that this function was only integrated into the online version in 2024, it is possible that it will also become available for the desktop version in the future.                                                                                                                                                                                                                                                                                                                                                   | că această funcție a fost integrată în varianta online abia în 2024, este posibil ca, în viitor, să existe și pentru varianta desktop.                                                                                                                                                                                                                                                                                                                                                            |
| Trados Online Editor includes commands which, in addition to the basic ones, such as stopping dictation or inserting different punctuation marks, can also be used to correct the text inside the segment through dedicated commands.                                                                                                                                                                                                                                                                              | Trados Online Editor include comenzi care, pe lângă cele de bază, cum ar fi oprirea dictării sau introducerea diferitelor semne de punctuație, pot fi folosite și pentru a corecta textul în interiorul segmentului, prin comenzi dedicate.                                                                                                                                                                                                                                                       |
| This cannot be done at the same level either with Dragon NaturallySpeaking or with Microsoft Voice Typing. Trados Online Editor has the ASR function specifically integrated into the translation process: you can move from one segment to another, lock segments, add tags, remove formatting, and the list goes on.                                                                                                                                                                                             | Acest lucru nu poate fi făcut la același nivel nici cu Dragon NaturallySpeaking, nici cu Microsoft Voice Typing. Trados Online Editor are funcția ASR special integrată pentru procesul de traducere: se poate trece de la un segment la altul, se pot bloca segmentele, se pot adăuga taguri, se poate șterge formatarea și lista continuă.                                                                                                                                                      |
| As I mentioned, there is hope that a plugin or broader desktop integration will also appear, especially since there is already a Text-to-Speech plugin, that is, the reverse version of Speech-to-Text. With it, we can listen directly to what we have translated and correct ourselves in this way; instead of reading every segment ourselves, we can listen to it as if it were a podcast. I did not present this option now because this time I wanted to focus on Speech-to-Text rather than Text-to-Speech. | După cum v-am spus, există speranța să apară și un plugin sau o integrare mai amplă pentru desktop, mai ales că există deja pluginul pentru Text-to-Speech, adică varianta inversă a Speech-to-Text. Cu acesta putem asculta direct ceea ce am tradus și ne putem corecta în acest mod; în loc să citim noi fiecare segment, îl putem asculta ca pe un podcast. Nu am prezentat această opțiune acum, pentru că de data aceasta am vrut să mă concentrez pe Speech-to-Text, nu pe Text-to-Speech. |
| And now I would really like to show you concretely how useful, from my point of view, the dictation function and the voice commands in Trados Online Editor can be, both for assistants and for translators.                                                                                                                                                                                                                                                                                                       | Și acum chiar aș vrea să vă arăt concret cât de utilă poate fi, din punctul meu de vedere, funcția de dictare și comenzile vocale din Trados Online Editor, atât pentru asistenți, cât și pentru traducători.                                                                                                                                                                                                                                                                                     |
| It is activated easily here on the right-hand side. We can see here on the left that it is working, that it is listening to us, and I would like to ask you to pay attention here, because this is where the voice commands that I dictate will appear in writing: “delete segment,” “lock segment,” “go to next one,” “confirm and go to next one,” “select next word,” “delete,” “confirm and go to next one”...                                                                                                 | Se activează ușor aici, în partea dreaptă. Vedem aici, în stânga, că funcționează, că ne ascultă, iar eu aș vrea să vă rog să fiți atenți aici, pentru că aici o să apară scrise comenzile vocale pe care eu le dictez: „șterge segmentul”, „blochează segmentul”, „mergi la următorul”, „confirmă și mergi la următorul”, „selectează următorul cuvânt”, „șterge”, „confirmă și mergi la următorul”...                                                                                           |
| I dictated with these pauses in speech, which are really more like thinking pauses, because I tried, to some extent, to simulate a real situation in which we read word by word, stop, do not speak very fluently, and translate while we are formulating our thoughts.                                                                                                                                                                                                                                            | Am dictat cu aceste pauze în vorbire, care sunt mai degrabă pauze de gândire, pentru că am încercat, într-o anumită măsură, să simulez o situație reală, în care citim cuvânt cu cuvânt, ne oprim, nu vorbim foarte cursiv și traducem pe măsură ce formulăm.                                                                                                                                                                                                                                     |

|                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                  |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| It can be used both for translating from scratch, when we dictate the whole segment, and for post-editing the segment that comes from translation memory or machine translation.                                                                                                                                                                 | Se poate folosi și pentru traducerea de la zero, când dictăm întregul segment, dar și pentru a post-edita segmentul care vine din memorie sau din machine translation.                                                                                                                                                           |
| Also, as you can see, I used the keyboard as well, in cases where it seemed more efficient to type than to dictate.                                                                                                                                                                                                                              | De asemenea, după cum puteți vedea, am folosit și tastatura, în cazul în care mi s-a părut mai eficient să tastez decât să dictez.                                                                                                                                                                                               |
| And that is basically it. The presentation could have been much longer, because there are many more solutions, but I wanted to briefly walk you through one solution for each area.                                                                                                                                                              | Și cam asta a fost. Prezentarea putea fi mult mai lungă, pentru că există mult mai multe soluții, însă eu am vrut să vă trec pe scurt prin câte o soluție pentru fiecare domeniu.                                                                                                                                                |
| In private life, in the car, in everyday situations when we want to send an email, in the professional environment through Dragon NaturallySpeaking, and in the translation environment through tools that we hope will gradually be integrated more fully into our workflow.                                                                    | În viața privată, în mașină, în viața personală când vrem să trimitem un e-mail, în mediul profesional prin intermediul Dragon NaturallySpeaking, iar în mediul traducerii prin instrumente care, sperăm, vor fi integrate treptat și mai bine în procesul nostru de lucru.                                                      |
| I hope you enjoyed this presentation and that you found it interesting and clear enough.                                                                                                                                                                                                                                                         | Sper că v-a plăcut această prezentare și sper că a fost destul de interesantă și clară.                                                                                                                                                                                                                                          |
| I am also going to send you an email with the information, the useful links and this video, in case you would like to watch it again. I also have a small request, because I am very interested in this topic and because I would like this ASR solution to be integrated into the translation process in the European Parliament in the future. | O să vă trimit și un e-mail cu informațiile, linkurile utile și acest video, în cazul în care vreți să îl revedeți. Aș avea și o rugămintă, pentru că mă interesează foarte mult acest subiect și pentru că mi-aș dori ca, în viitor, această soluție ASR să fie integrată și în procesul de traducere din Parlamentul European. |
| What do you think? Would you use this application? I would be very glad to receive your feedback. It is a short questionnaire, with questions rated from one to five, as well as multiple-choice and open-ended questions, none of which are mandatory.                                                                                          | Ce părere aveți? Ați folosi această aplicație? Mi-ar plăcea foarte mult să primesc feedback din partea dumneavoastră. Este vorba despre un scurt chestionar, cu întrebări evaluate de la unu la cinci, cu răspunsuri multiple sau deschise, care nu sunt obligatorii.                                                            |
| And, if you are interested, I have Dragon NaturallySpeaking installed on my Parliament laptop, or I can show you on my personal laptop the other ASR options that are not available on Parliament devices. Thank you very much, and if you have any questions, please do not hesitate.                                                           | Și, dacă sunteți interesați, eu am Dragon NaturallySpeaking instalat pe laptopul Parlamentului sau vă pot arăta pe laptopul meu personal și celelalte opțiuni ASR care nu sunt disponibile pe dispozitivele Parlamentului. Vă mulțumesc mult și, dacă aveți întrebări, vă rog să nu ezitați.                                     |

## **Appendix D - Follow-up questionnaire (EP RO Unit, DG TRAD)**

### **Questionnaire**

[https://docs.google.com/forms/d/e/1FAIpQLSdtJWDteUFz3CSycEKb3Hadc4zPkfWKu46w\\_h4V97h4n2Idsg/viewform?usp=dialog](https://docs.google.com/forms/d/e/1FAIpQLSdtJWDteUFz3CSycEKb3Hadc4zPkfWKu46w_h4V97h4n2Idsg/viewform?usp=dialog)

### **Questionnaire results**

Romanian original:

[https://docs.google.com/spreadsheets/d/1vVJJXyVSiCQ\\_cbvZMJ4p9l3DrFuJ9rey/edit?usp=sharing&ouid=111344779374230123077&rtpof=true&sd=true](https://docs.google.com/spreadsheets/d/1vVJJXyVSiCQ_cbvZMJ4p9l3DrFuJ9rey/edit?usp=sharing&ouid=111344779374230123077&rtpof=true&sd=true)

English translation:

<https://docs.google.com/spreadsheets/d/1BkQX7p21EAGALypeX87de22IZrcBzDrP/edit?usp=sharing&ouid=111344779374230123077&rtpof=true&sd=true>

## ACKNOWLEDGEMENTS

This thesis could not have been completed without the support of several people to whom I would like to express my sincere gratitude. First, I would like to thank my supervisor, Univ.-Prof. Dragoş Ioan Ciobanu, for his guidance, encouragement and valuable feedback throughout the development of this thesis.

I am also grateful to my colleagues at the European Parliament's Directorate-General for Translation (DG TRAD) for their openness and support. My thanks go to the Head of the Speech-to-Text Unit for sharing valuable institutional insights, to the anonymous Romanian proofreader who contributed the Verba user perspective, and to my colleagues in the Romanian Unit, whose participation in the presentation, follow-up try-out and questionnaire made the empirical part of this study possible.

Conducting this research within the professional environment in which these questions arise gave me valuable access to relevant people, information and discussions, for which I am very thankful.