Trainee Translators' Perceptions of the Role of Pronunciation and Speech Technologies in the Technology-Driven Translation Profession

ABSTRACT

We live in a world of rapid technological advances which constantly affect the work of professional translators. Suitable training is therefore required for future translators to be able to compete on the translation market. With the rise of translation technologies, new ideas have been put forward on how to make translators faster and more efficient. Among the technologies that future translators may not be adequately familiar with are speech recognition tools; these enable translators to dictate their sight translation and have it typed out, allowing more time to focus on the content. However, as with all digital tools, the quality of input is important; a question thus arises on the role pronunciation assumes in such work. The present study aimed to establish how much awareness there is amongst the trainee translators of the possibilities afforded by speech technologies and to explore their perceptions of the role played by pronunciation.

Keywords: translator training; pronunciation; speech recognition tools; trainee translators' perceptions; the future of translation work

Bodoči prevajalci o vlogi izgovarjave in govornih tehnologij v sodobnem prevajalskem poklicu

POVZETEK

Živimo v času vse hitrejšega tehnološkega razvoja, v kar je nenehno vpeto tudi delo profesionalnih prevajalcev. V luči tega je nujno sprotno prilagajanje izobraževanja bodočih prevajalcev, da bodo primerno usposobljeni in bodo konkurenčni na prevajalskem trgu. S porastom sodobnih prevajalskih tehnologij se pojavljajo ideje o tem, kako bi lahko bili prevajalci pri svojem delu hitrejši in učinkovitejši. Eden od tehnoloških pripomočkov, ki bi k temu lahko pripomogel, a ga bodoči prevajalci premalo poznajo, so govorne tehnologije. S pomočjo prevajanja na vpogled prevajalec lahko besedilo narekuje: s tem se izogne tipkanju, in se bolj osredotoča na vsebino. A kot pri vseh digitalnih orodjih je pomembna kakovost vnosa podatkov, zato se poraja vprašanje, kakšno vlogo igra pri tem izgovarjava. V pričujoči študiji smo želeli raziskati, v kolikšni meri se bodoči prevajalci zavedajo možnosti, ki jih ponujajo govorne tehnologije, in ali imajo predstavo o vlogi, ki jo pri tem igra izgovarjava.

Ključne besede: poučevanje prevajalcev; izgovarjava; govorne tehnologije za razpoznavo govora; zavedanje bodočih prevajalcev; prevajalsko delo v prihodnosti

1   Introduction

The impact of new technologies on translation work over the last few decades has significantly changed the way people perceive the work of professional translators. The usual translator's workstation or translator's workbench no longer involves working only with computers and computer-assisted (CAT) tools, but may, under certain conditions, also involve working with machine translation (MT) and speech recognition technologies. According to a Stanford study (cf. Weiner 2016[1]) speaking is much faster than typing on a touchscreen, while typing on a computer keyboard is seemingly easier and faster. However, even a few years ago speech recognition software was criticised due to its error-prone performance which inevitably lead to spending too much time correcting the mistakes. It therefore seemed reasonable to assume that professionals who use a keyboard as part of their daily routine, translators included, would not be inclined to integrate into their work technologies which actually slow them down. However, a lot has changed since then: Nuance has produced Dragon Speech Recognition software, one of the leading speech recognition technologies, and claims that it is now able to transcribe up to 160 words per minute, which is also about three times faster than typing, with an enviable 99% recognition accuracy (cf. Dragon NaturallySpeaking[2]). This suggests speech technologies are now much more effective, and can perhaps make translation work more efficient. Moreover, any technological advantage is worth exploring to ensure that professional translators remain competitive on the translation market.

With the swift rise of digital innovations and artificial intelligence (AI), significant endeavours will constantly, and increasingly so, be put into speech technologies for translation undertakings, at least for fairly basic communication purposes and simple translation tasks, with the aim to establish basic contact and ease communication for those who do not speak a particular language. Students might already be aware of the possibilities afforded by virtual AI speech assistants such as Amazon's Alexa, Microsoft's Cortana, Google's Assistant or Apple's Siri, and might have tried using such services. Large brands are all investing heavily into voice technologies, and they are associated with a growing number of applications (cf. for more details on virtual assistants see Moren 2018). Armour (2018) reports on the data provided by Adobe Analytics, which indicates that "71% of owners of smart speakers like Amazon Echo and Google Home use voice assistants at least daily" [...] with "44% using them multiple times a day" while "[o]ver 76% of smart speaker owners increased their usage of voice assistants in the last year". Armour (2018) also quotes Steve Rabuchin, VP of Amazon Alexa, who stated that the vision they have for their customers is to "be able to access Alexa whenever and wherever they want. This means customers may

---

[1]   Cf. https://www.popularmechanics.com/technology/a22684/phone-dictation-typing-speed/.
[2]   Compare with data provided by Nuance at https://www.nuance.com/dragon/industry/education-solutions.html.

be able to talk to their cars, refrigerators, thermostats, lamps and all kinds of devices in and outside their homes". Armour (2018) believes that "voice is the future of how brands will interact with their customers". These virtual assistants are all monolingual, however, and do not engage in multi-lingual communication. Even so, "[t]o build a robust speech recognition experience, the artificial intelligence behind it has to become better at handling challenges such as accents and background noise. And as consumers are becoming increasingly more comfortable and reliant upon using voice to talk to their phones, cars, smart home devices, etc., voice will become a primary interface to the digital world and with it" (Armour 2018).

Virtual assistants no longer work only with English[3]; Cortana, for example, is currently also available in Chinese, French, German, Italian, Japanese, Portuguese and Spanish versions, making these voice technologies increasingly accessible to a much wider audience[4]. Even regular dictation services available to Windows and Mac users have the option of choosing between language varieties, with American, Australian, British or Canadian English, for example, already embedded while, depending on the tool, other varieties can easily be downloaded from the Internet. However, more time may be required to have languages of lesser diffusion[5] successfully integrated into existing systems. Slovene is a language spoken by only about two million people, and thus is less likely to be automatically added to other major language options. However, there are some speech recognition tools available, such as Voice Notepad, which already have Slovene embedded, and the dictation performance is relatively accurate. This is in contrast to the Google Translate dictation option, as the quality of translation is often still highly questionable and the final output more frequently than not inadequate and unusable. There is even a virtual AI assistant SecondEGO, designed by Amebis[6], and several other systems available for Slovene, which were originally created on the basis of large corpora and other language resources[7], such as the speech-to-speech communicator VoiceTRAN[8] or eBralec[9] (eReader): the direction, however, is speech-to-speech or written to spoken rather than spoken to written, which would be most suitable for translators. Moreover, these technologies are only available commercially or for research purposes (cf. Sepesy Maučec et al. 2009; Donaj and Kačič 2012; Žgank and Sepesy Maučec 2010; Žgank, Verdonik, and Sepesy Maučec 2016, to name just a few), while their non-commercial availability is still a matter for the future.

---

[3]   Other languages are also gaining ground on the Internet (cf. Internet World Stats 2017).

[4]   For more on English and its relative share online see Holly Young's article available at http://labs.theguardian.com/digital-language-divide/ and Laura Gonzales' article available at http://uxpamagazine.org/improving-digital-translation/.

[5]   Slovene included (cf. Pokorn 2005; Hirci 2012).

[6]   Cf. https://www.amebis.si/novice/npi-2015.

[7]   For more on Slovene in the digital age see Rehm and Uszkoreit (2012).

[8]   Cf. http://www.alpineon.si/voicetran/slovensko/html/index.html.

[9]   Cf. https://ebralec.si/?jezik=sl.

Still, none of these technologies are directly applicable to regular translation work as they are aimed at the general public to ease their daily routines. None of the virtual assistants are applicable to ease the tedious task of typing which has to be regularly undertaken by translators; translators thus need more specialised translation tools to facilitate their work (cf. Cronin 2013). One option that could possibly aid their daily routines and reduce the need for constant typing is dictation. Combined with sight translation it could change the way translation is habitually performed. It might thus be worth investigating the usability of speech-to-text technologies in translator training, foregrounding the time-efficiency ratio in particular. The awareness of trainee translators of the role of pronunciation and their familiarity with speech recognition technologies deserve research attention, in order to establish whether the application of such technologies could be motivating and beneficial for future translators.

2   Literature Review

Professional translation work is usually associated with the written output. However, the spoken modality should not be neglected in today's information society and its digital world, so heavily imbued with multimodality. It is therefore worth exploring the issues in translator training that address these modalities, spoken included, especially since – within the scope of interpreter training – Shlesinger (1995, 193–214) already maintained that "one modality can teach us about the constraints, conventions and norms of the other". This suggests that sight translation, a bridge between the oral and written mode of translation (cf. Agrifoglio 2004), should perhaps play a more prominent role not only in professional translation, but also in translation pedagogy.

So far, sight translation has been recognised as relevant in interpreting studies and interpreting pedagogy (cf. Agrifoglio 2004; Angelelli 1999; Li 2014; Gile [1995] 2009; Gonzalez, Vásquez, and Mikkelson 2012; Jimenez Ivars 2008; Lambert 2004; Mikkelson 1994; Moser-Mercer 1995; Pöchhacker 2004, 2010; Riccardi 2002; Schlesinger 1995; Song 2010; Viaggio 1995; Viezzi 1990; Weber 1990). Although there is still a fairly small body of literature focusing on the advantages of sight translation for written translation (cf. Baxter 2016; Dragsted and Hansen 2009; Dragsted, Hansen, and Sørensen 2009; Dragsted, Mees, and Hansen 2011; Gorszczyńska 2010; Mees et al. 2013), a recent study has shown (cf. Hirci, Mikolič Južnič, and Pisanski Peterlin forthcoming) that engaging in sight translation for the purposes of written translation can result in creative, novel translation solutions, which gives an added value to the translation process and can make the entire process of translating much faster and more efficient. Some scholars have already explored the application of dictation in sight translation and foregrounded its benefits for translation work in terms of time efficiency (cf. Biela-Wolonciej 2007). Possible advantages were also reported by Dragsted, Mees, and Hansen (2011), who compared written and sight translation output with and without speech recognition software.

They concluded that with additional training and better familiarity with speech recognition tools, "greater time savings and higher quality are likely to be achieved as technical obstacles are either reduced or overcome" (Dragsted, Mees, and Hansen 2011, 26). Baxter (2016) also investigated the application of sight translation skills to written translation combined with speech recognition; although there were no considerable time differences for the two studied groups, idiomaticity was enhanced, suggesting that combining sight translation with speech recognition "improves the spontaneity of the final text, thereby producing a more natural-sounding translation than the traditional W2W[10] method" (Baxter, 2016, 14). However, the most interdisciplinary approach was adopted in a study by Mees et al. (2013) where close collaboration among phoneticians, translators and interpreters yielded sound grounds for further interdisciplinary cooperation, proving that speech recognition technologies[11] can be successfully applied in translator training.

In Slovenia, no study has been carried out on having speech recognition technology fully integrated into translation work, focusing on a hybrid which "involves crossing borders between translation and interpreting since the translation is produced orally, as in interpreting, but is visible on the screen, as in translation" (Mees et al. 2013, 141). There is an introductory course on English phonetics and phonology for translators offered in year one of the undergraduate programme at the Department of Translation Studies in the University of Ljubljana to help students improve their pronunciation. As the advances in speech-to-text technology are relatively recent, students enrolled in the course may not be familiar with the relevance of pronunciation skills in technological applications, and may perceive pronunciation to be more important for interpreters than translators. Yet this issue is particularly relevant for those who may wish to use software which is heavily reliant on one's pronunciation. As Nuance is claiming a 99% accuracy for its software, it needs to be acknowledged that such accuracy is only possible if one's pronunciation is also highly accurate, otherwise the success rate of speech recognition is much lower. Near-native and intelligible pronunciation is required for the dictation systems to work well, at least for the time being, otherwise the rate of mistakes due to mispronunciation is too great to have such tools considered effective. However, so far the potential relevance of pronunciation skills for the trainee translators' work in the translation modules offered later as part of the graduate programme in Translation/Interpreting has not yet been addressed, as none of the specialised translation courses involve working with speech recognition technologies. As there are built-in dictation options available on computers (both for Windows and Mac users) that enable working with English, translation modules focusing on translation from L1 to L2 could possibly benefit

---

10   W2W means written to written translation.
11   For more details on speech recognition technology see Jurafsky and Martin (2000).

from integrating this technology into their regular translation instruction. In their study, Mees et al. (2013) also report on working into L2 (cf. studies by Dragsted and Hansen 2009; Dragsted, Mees, and Hansen 2011). In Denmark, and the rest of Scandinavia, where, according to Phillipson (2003, 96) there are "good grounds for referring to English as *a second language* rather than a foreign language", working into L2 is not perceived as unusual. Both Danish as well as Slovene are comparable in this respect, as they can both be considered as languages of lesser diffusion, so L1 to L2 translation (cf. Pokorn 2005; Hirci 2012) is not uncommon in Slovenia either. In fact, children in Slovenia start learning English as part of their primary school curriculum at the age of six. Films and TV shows are regularly subtitled rather than dubbed, and Slovene translators work into both directions, L2 to L1 as well as L1 to L2. Many professional translators in Slovenia find themselves in a position where they are required to undertake translation into L2, English in particular, on a regular basis, since there is a serious shortage of native English speakers working with Slovene. Thus training is necessary in the L2 direction and is offered as part of the translator training curriculum at the Department of Translation Studies in the University of Ljubljana.

2.1 Future Prospects – More Work with Speech Recognition Systems?

So far, no research has been undertaken in Slovenia to explore working with speech recognition systems focusing on time efficiency in translation. However, a study was carried out on the possible benefits of applying speech recognition technologies in the pronunciation training of non-native speakers of English. Šuštaršič (2005, 87) investigated some software packages to explore their "usability within an English phonetics curriculum for EFL learners at the university level" that can be applied to pronunciation training. Šuštaršič (2005, 93–97) suggested that "speech recognition can be applied in phonetics (or more precisely, in pronunciation) teaching, and that a number of aspects of articulatory and auditory phonetic principles can be observed in the way that speech recognition programs transfer (or fail to transfer) the received speech signals into written form." He pointed out that "using any speech recognition program with English pronunciation students has several other justifications. Firstly, the program needs to be trained to one's voice, which requires a great deal of loud reading. […] The basic rule is: the more you train the program (i.e. the more you read), the higher will be the accuracy of recognition, and thus the usefulness of the program for any practical task." Šuštaršič (2005, 98) also suggested that students can be encouraged to record their own speech and apply a speech recognition programme to convert it into a written text, an idea which in itself is closely related to sight translation from Slovene into English. Šuštaršič (2005) reported working with commercial speech recognition technologies such as Via Voice and Dragon's NaturallySpeaking, which, however, are not freely available. A cost-free option

nowadays is to simply activate the automatically built-in dictation option on the computer (either for Windows or Mac users), as it comes at no additional price, and explore its usability before obtaining some more sophisticated commercial software.

Drawing on Mees et al. (2013) and Šuštaršič (2005), a study was thus conceived to explore the possible benefits of using speech technologies in translator training for two reasons:

- to improve trainee translators' pronunciation,
- to use speech instead of typing to speed up the process of translation.

3 Study Design and Methodology

The present study was designed to explore the trainee translators' perceptions of the role of English pronunciation, as well as their familiarity with speech recognition tools, to establish whether or not it might be viable to introduce such technologies into translator training at the University of Ljubljana.

3.1 Methodology and Participants

An online questionnaire was designed for the purposes of the present study to foreground the perceptions of both undergraduate and graduate trainee translators studying at the Department of Translation Studies at the Faculty of Arts, University of Ljubljana, in the academic year of 2018/2019.

3.2 Data Collection

The questionnaire was made available online for 18 days, between 4 January 2019 and 22 January 2019, with a total of 94 participants taking part in the study. The questionnaire, designed using the online Google Forms survey mode, consists of 18 questions. The first part of the survey aims to collect general information about the participants, eliciting data on their age, gender and year of study. The second part of the questionnaire explores the participants' perceptions and self-awareness of their own pronunciation and their familiarity with the existing speech-to-text technologies that might prove to be useful in their future profession.

The trainee translators were asked to respond to several statements referring to their perceptions of the role pronunciation in English and their aspirations to improve it (i.e. a total of nine questions corresponding to yes/no answers, and four statements using a five-point Likert-type scale, ranging from "totally unmotivated" = 1 to "extremely motivated" = 5 related to the participants' motivation to have good pronunciation of English, from "the least important" = 1 to "the most important" = 5 on how important they find pronunciation in relation to other language skills, from "extremely poor" = 1 to "excellent" = 5 on how they would rate their own pronunciation at the time of

filling out the questionnaire, and finally from "do not aspire to this at all" = 1 to "aspire to this 100%" = 5 on how much they aspire to have a near-native pronunciation of English).

Additional information on the existing speech recognition tools and the students' experience with the application of these technologies to their work was elicited using a number of multiple choice questions. The participants were also encouraged to provide additional comments on the possible benefits of using speech-to-text technologies in the final section of the questionnaire.

4   Results and Discussion

This section reports on the results of the questionnaire completed by the participants of the study. First, general demographic information on the participants is provided, followed by the data related to their pronunciation and awareness of speech recognition technologies. Due to the limited scope of this paper only those results that directly address the topic are discussed in detail.

4.1 General Information on the Participants

The study involved 94 participants, of whom all completed the questionnaire in full. All of the participants are either undergraduate BA students of Interlingual Mediation, or graduate MA students of Translation/Interpreting in the University of Ljubljana (cf. Figure 1). Of the 94 participants, 76 were female and 18 were male, and all were aged between 17 and 26 (average 21).

Most participants (41, i.e. 43.6%) are enrolled in year 1 of the BA in Interlingual Mediation, with 14 (14.9%) respondents from year 2 of the BA in Interlingual Mediation, and 16 (17%) respondents from year 3 of the BA in Interlingual Mediation (cf. Figure 1). At the graduate level, there were 15 (16%) participants from MA I in Translation, three (3.2%) from MA I in Interpreting, and five (5.3%) from MA II in Translation (there is no MA II in Interpreting available for this academic year).

94 responses

FIGURE 1. Participants in the study (N=94).

4.2 Specific Information on Pronunciation

Importance to speak English well

As evident from the results of the questionnaire, all of the participants believe that it is important to speak English well to make a good impression on their clients and employers, and all but one believe the same is important to be a successful interpreter, while 88 out of 94 participants (i.e. 93.6%) were of the opinion that this is also important for translators (cf. Hirci 2017). In addition, 90 (95.7%) respondents think that it is important to speak well to sound professional, and 83 (88.3%) to be able to use speech recognition tools more easily.

Significance of speaking English well

The participants seem to have rather diverse views on what speaking English well actually means. Most of the participants, i.e. 89 (94.7%), agreed that this meant having pronunciation which is intelligible and easy-to-understand, with 65 (69.1%) believing it meant speaking with an accent which is close to standard varieties of English. Fewer than half of the respondents in all (45 or 47.9%) believe that this meant having a native-like pronunciation.

Motivation to have a good pronunciation of English

The questionnaire yielded an insight into the participants' motivation with regard to having good pronunciation: the results show that over half of the participants (52 or 55.3%) are extremely motivated and an additional 30 (31.9%) are very motivated to have a good pronunciation of English (a mean score[12] of 4.4, cf. Table 1), which confirms that the respondents regard having good pronunciation in English as essential for their future profession.

Table 1. Mean scores for pronunciation.

Perceptions about pronunciation	Mean score
Motivation to have a good pronunciation of English	4.4
Importance of pronunciation compared to other language skills	3.8
Assessment of own pronunciation	3.5
Aspirations to improve their pronunciation	4.3

When asked about how important they find pronunciation compared to other language skills, the participants showed considerable agreement that pronunciation skills are quite important (a mean score of 3.8).

---

[12]   The central tendency for each Likert-type statement was summarised using the mean score.

The participants' replies furthermore revealed that they tend to aspire to have English pronunciation which is intelligible yet close to one of the English standards. They deemed their own pronunciation at the time of filling out the questionnaire as only "good" or "fairly good", while only two participants considered it "excellent". Three participants even believed their pronunciation was "extremely poor" or "rather poor" (mean score 3.5). The responses revealed that over half (54.3%) of the participants have extremely high aspirations to improve their pronunciation, and an additional 29.8% of the participants have high aspirations to improve it (mean score 4.3).

These results are quite valuable, as they reveal that most participants are aware of the significance of having a good pronunciation of English. Whether they see a correlation with speech-to-text technologies, however, is yet to be explored. As clear, accurate and intelligible pronunciation is required to have speech recognition systems work well, at least for the time being, improving non-native English pronunciation is undoubtedly worth investing time and effort into if we also wish to gain from the advantages afforded by such technologies.

4.3 Specific Information on Speech Recognition Technologies

We wished to establish if the respondents were aware of the differences in speed as related to speech and typing. According to Nuance's Dragon speech recognition software, speaking is three times faster than typing. Most respondents of this study, i.e. 45 (47.9%), believed that speech was two times faster than typing, while 37 (39.4%) participants in fact responded that it was actually three times faster. Only two participants were of the opinion that speaking was slower than typing, three assumed that it was four times faster, while another four responded that these two activities were both of equal speed (cf. Figure 2, where responses are provided as option Other, after the option 4x faster).

94 responses

- speaking is slower than typing
- equal speed
- 2x faster
- 3x faster
- 4x faster
- it depends on the speed of the speech and typing
- faster (hard to quantify how much, depends on the typing skill)
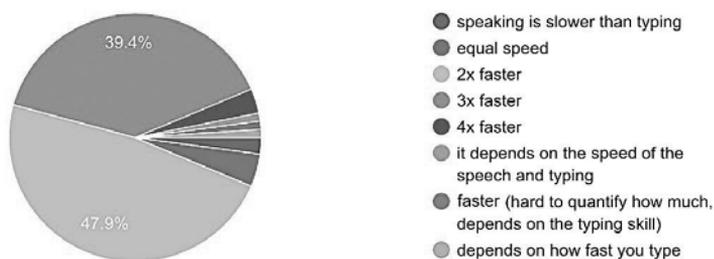- depends on how fast you type

FIGURE 2. Speed of speaking v typing (N=94).

It was no surprise to see that almost half of the participants (44.7%) responded that they have already used the built-in dictation software on their smartphones; nevertheless, the number is much lower for computers, where only 11 of participants out of 94 reported using this technology. Interestingly enough, 28 of the participants reported that their dictation was successful, or at least sometimes or to some extent. It is fair to assume that with more accurate pronunciation of English the perception of the success rate would most likely be even higher. Some participants also pointed out that they used the dictation option only on their smartphones, without ever realising that this was also possible on their computers.

In all, 70 (74.5%) of the participants responded that they would consider using dictation in their translation work; even more, i.e. 84 (89.4%) believed that it would be useful to work with speech recognition tools as part of their translator training at the university. In additional, individual comments, the participants provided a number of reasons why they assumed it would be useful to work with speech recognition tools as part of translator training (cf. Figure 3).

P14: *"It could improve the student's pronunciation skills and, more importantly, the proper flow of speech."*
P8: *"Speech recognition tools are great for improving ones pronunciation and I think we should focus on that and phonetics in general more thoroughly."*
P4: *"I believe that students should be familiar with any translation- or language-related technology. This can be useful in their careers."*
P53: *"I think that such thing as a speech recognition tool would help me a lot with my poor pronunciation."*
P16: *"Working with these tools would improve our pronounciation."*
P15: *"I think we would be able to translate everything faster. And we would also practice our pronunciation and expand our vocabulary, because when we say something outloud, we remember it faster."*
P18: *"So that we learn different approaches to translating and figure out for ourselves which best suits us. Also I think it is less time consuming than typing and prevents you from making spelling mistakes"*
P23: *"It's a tool that is becoming increasingly popular and it could potentially make future work easier."*
P29: *"Knowledge of new technologies is always useful, the more you know the more you can learn, new skills can easily improve our employability, variation of skills is important for adapting to the market"*
P19: *"The more education we get - conected to our studies and technology connected to languages – the better."*
P20: *"Speech recognition tools are developing and becoming a bigger part of our everyday life".*

P42: "It would improve out studying and it would be a variation of "teaching" that is not often used."
P49: "Because any aspect of the translation work that we are presented is welcome and useful. Anything that we learn might come in handy and we are better because of each of those experiences."
P56: "So that we learn different techniques and figure out which approach best suits us."
P59: "The advancement of technology will impose these tools sooner or later and it would be best if the new generations of translators and interpreters had mandatory training with them."

FIGURE 3. Comments provided by the participants.[13]

These comments show that there is already some degree of awareness amongst the trainee translator population of the possible advantages associated with the integration of speech recognition technologies into translator training.

The results of the questionnaire related to the various types of speech recognition software that the participants might have heard of are specified in Figure 4. The most frequently recognised speech recognition technologies were Windows Speech Recognition (60), Apple's dictation (49) and Google Docs Voice Typing (48), followed by IBM's Speech to Text (38), Amazon's Transcribe (27) and Speechnotes (25). The other speech-to-text tools (such as Via Voice, Dragon NaturallySpeaking or Voice Finger) were much less frequently recognised, while only one participant in this study had heard of Braina Pro.
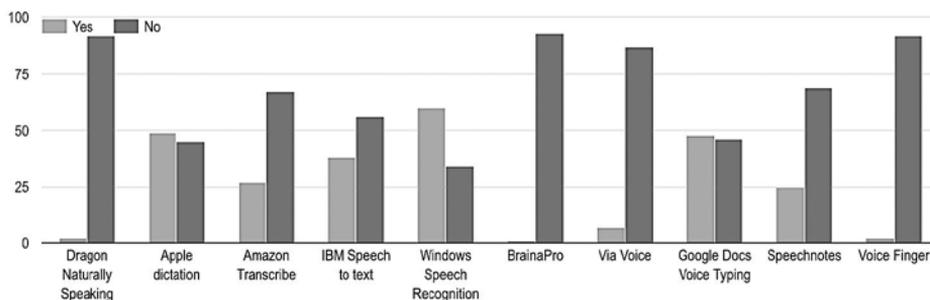

FIGURE 4. Familiarity with speech recognition technologies (N=94).

In addition, only three other online speech recognition tools were mentioned by the participants who were offered an option to list any other speech recognition technologies of which they might be aware: one of the participants noted using Google Keep, while another participant had not only heard of but has tried Voice Notepad for Slovene (they reported, however, that their dictation work was not

---

[13]   All comments by the participants are provided in their original form, verbatim, with spelling mistakes and other errors left unchanged.

highly successful). It is interesting to note that 32 (i.e. 34%) of the participants responded that they have already tried using some of the tools mentioned, selecting mainly Apple Dictation, Google Docs Voice Typing and Windows Speech Recognition (only two participants selected Speechnotes, while only one mentioned IBM's Speech to text and another one Via Voice, cf. Figure 4). Most of the participants (62, or 66%) learnt about these speech tools online, by themselves, and only five (i.e. 5.3%) at the university.

P13: "It's a faster way of writing down what you need to translate and possibly a more fun and/or interesting way of translation."
P15: "Good speech recognition tools could help us learn proper pronunciation."
P20: "Because they could save quite a lot of time and work for translators (there would be also be no typos in the text etc)."
P23: "An extra aid one might find useful (like a dictionary or a thesaurus)."
P29: "They could be useful for translating things that need to be transcribed anyway, like speeches, directly, or just as an alternative to typing."
P40: "For transcribing and general text formation – the limits of one's typing skill can cause the occurence of getting lost in thought while typing and forgetting what you were about to say. In speech it happens less often"
P44: "It could be helpful if one has to translate videos or with subtitling."
P48: "These tools can facilitate the translation of audio documents"
P58: "They could replace typing, which can be time-consuming and tiring."
P65: "We could see where the problem with our speech is."
P67: "Because it is useful knowing tools that can make the translation work easier. This presents us with what the translation work is like and prepares us for it."
P68: "It is faster, so they can earn more money in a shorter period of time and thus have more free time. :)"
P70: "These tools could mean that translators would finish their work faster. Some may speak faster than they type so it could improve their working conditions."
P75: "To facilitate transcribing spoken language, could be useful for making subtitles"
P81: "Since speech recognition tools are very accurate nowadays, I believe it would save a lot of time."

FIGURE 5. Participants' comments on the usefulness of speech recognition tools for translators (also verbatim).

Judging from the comments provided in the questionnaire, some participants are also aware of the drawbacks of the current speech recognition technologies and their reliability:

> P7: "*Translators could work faster, but the speech recognition tools would need to be very good, especially when it comes to punctuation. Going over a text two or more times to correct punctuation that was wrongly placed by speech recognition tools is very time consuming.*"
>
> P37: "*They might be useful for cases, when translators have to write subtitles e.g. a speech or movie, and are not sure about what a person is saying. However the speech recognition tools are not yet reliable enough to be completely sure of whether their result is correct.*"

As it can be observed from the participants' comments, the predominant idea revolved around the opinion that speech is "faster than typing", and that the application of speech technologies could make translators more efficient. Some students are well aware of the current situation in the ever-evolving digital world, recognizing that "the use of speech recognition today is growing and many people use it on their phones (Siri) or have devices (Alexa) that help them with everyday tasks." (P34)

All this suggests it might be worth raising the awareness of the trainee translator population about the existence of such tools, and possibly even integrate speech recognition technologies into translator training. This could be achieved in several ways: either by implementing information on speech recognition technologies into the already existing technology-related courses, or by introducing it as part of a new course focusing on this particular topic with hands-on training within L1 to L2 translation modules.

Some studies have already shown (cf. Mees et al. 2013; Désilets et al. 2008) that the implementation of speech technologies into translation work is something that could possibly be better addressed in the future. There are also interesting pedagogical implications of this: if dictation may soon become an increasingly dominant mode of communication, it is important to gain an in-depth insight into the aspects of pronunciation that would be particularly relevant in translator training.

5 Conclusion

The present study explored the perceptions of trainee translators studying at the Department of Translation Studies in the University of Ljubljana on pronunciation and speech technologies. The results of the study offer good grounds for a more prominent role to be assigned to both pronunciation instruction and speech technologies in translator training. The study yielded results showing that an overwhelming majority of trainee translators (just under 90%) believe that having good pronunciation of English is important for their profession (cf. Hirci 2017), while over 80% also have aspirations to improve their pronunciation. In addition, the results show that all the participants believe it is important to speak English well to make a good impression on clients and employers; all but one find this important for interpreters, while 93.6% also find it important for translators. Moreover,

95.7% respondents stated that it is important to speak well to sound professional, and 88.3% believe this is important to be able to use speech recognition tools more easily.

These results suggest that equipping trainee translators with pronunciation skills for speech recognition technologies is of relevance and would most likely be embraced by the students. This is in line with the study by Mees et al. (2013, 149), whose retrospective interviews revealed that "a number of students feel that they have become more aware of their pronunciation problems in the course of training the SR [speech recognition] program". Their study also revealed that speech recognition "provides a potentially useful supplement to written translation, or indeed an alternative to it" (Mees et al. 2013, 140–42). The immediate time-efficiency aspect is therefore yet another reason why speech recognition technologies could be applied in translator training: a new modality could also enhance the learning experience in the translation classroom. As some participants of this study have observed, "Time is valuable. Every second saved from sitting in front of a screen and keyboard is warmly welcome" (P41) or "It is faster, so they can earn more money in a shorter period of time and thus have more free time. :) (P68)."

With the increasingly rapid advances in voice activated technologies, translator trainers should seize the opportunity to retain tech-savvy students' interest and channel it into their regular coursework. Staying ahead is vital to remaining competitive; having that special 'edge' might be a deciding factor in having trainee translators turn into successful players on the professional translation market. Thus aiming to have good pronunciation and speak English well enough to be able to work with speech recognition technologies could prove to have added value for translators' professional careers.

References

Agrifoglio, Marjorie. 2004. "Sight Translation and Interpreting: A Comparative Analysis of Constraints and Failure." *Interpreting. International Journal of Research and Practice in Interpreting* 6 (1): 43–67. http://dx.doi.org/10.1075/intp.6.1.05agr.

Angelelli, Claudia. V. 1999. "The Role of Reading in Sight Translation." *The ATA Chronicle (Translation Journal of the American Association of Translators)* 28 (5): 27–30.

Armour, Britt. 2018. "7 Key Predictions for the Future of Voice Assistants and AI." Accessed January 15, 2019. https://clearbridgemobile.com/author/britt_clrbridge/.

Baxter, Neal. R. 2016. "Exploring the Effects of Computerised Sight Translation on Written Translation Speed and Quality." *Perspectives* 1: 1–18. https://doi.org/10.1080/0907676X.2016.1241287.

Biela-Wolonciej, Aleksandra. 2007. "A-vista: New Challenges for Tailor-Made Translation Types on the Example of Recorded Sight Translation." *Kalbotyra* 57 (3): 30–39.

Cronin, Michael. 2013. *Translation in the Digital Age*. London: Routledge.

Désilets, Alain, Marta Stojanovic, Jean-François Lapointe, Rick Rose, and Aarthi Reddy. 2008. "Evaluating Productivity Gains of Hybrid ASR-MT Systems for Translation Dictation." In *IWSLT 2008, International Workshop on Spoken Language Translation, 20–21 October 2008, Waikiki, Hawai'i, USA*, Waikiki, Hawai'i, 158–65.

Donaj, Gregor, and Zdravko Kačič. 2012. "Širjenje slovarja in dvoprehodni algoritem v razpoznavalniku tekočega govora UMB Broadcast News." In *Proceedings of the Eighth Language Technologies Conference, October 8th-12th, 2012,* Ljubljana, Slovenija, 48–51. Ljubljana: Institut Jožef Stefan.

Dragsted, Barbara, and Inge G. Hansen. 2009. "Exploring Translation and Interpreting Hybrids. The Case of Sight Translation." *Meta: Journal des traducteurs/Meta: Translators' Journal* 54 (3): 588–604. https://doi.org/10.7202/038317ar.

Dragsted, Barbara, Inge G. Hansen, and Henrik S. Sørensen. 2009. "Experts Exposed." In *Methodology, Technology and Innovation in Translation Process Research* (Copenhagen Studies in Language 38), edited by Inger M. Mees, Fabio Alves and Susanne Göpferich, 293–317. Copenhagen: Samfundslitteratur.

Dragsted, Barbara, Inger M. Mees, and Inge G. Hansen. 2011. "Speaking Your Translation: Students' First Encounter with Speech Recognition Technology." *Translation & Interpreting* 3 (1): 10–43.

Gile, Daniel. [1995] 2009. *Basic Concepts and Models for Interpreter and Translator Training.* Amsterdam/Philadelphia: John Benjamins Publishing.

Gonzalez, Roseann D., Victoria F. Vásquez, and Holly Mikkelson. 2012. *Fundamentals of Court Interpretation: Theory, Policy and Practice.* 2nd ed. Durham, North Carolina: Carolina Academic Press.

Gonzales, Laura. 2017. "Improving Digital Translating: Research Findings from Multilingual Communicators." *User Experience Magazine* 17 (5). http://uxpamagazine.org/improving-digital-translation/.

Gorszczyńska, Paula. 2010. "The Potential of Sight Translation to Optimize Written Translation: The Example of the English-Polish Language Pair." In *Translation Effects. Selected Papers of the CETRA Research Seminar in Translation Studies 2009,* edited by Omid Azadibougar, 1–12. Leuven: KU Leuven. https://www.arts.kuleuven.be/cetra/papers/files/paula-gorszczynska-the-potential-of-sight.pdf.

Jimenez Ivars, Maria. 2008. "Sight Translation and Written Translation. A Comparative Analysis of Causes of Problems, Strategies and Translation Errors within the PACTE Translation Competence Model." *FORUM. International Journal of Interpretation and Translation* 6 (2): 79–104.

Hirci, Nataša. 2012. "Electronic Reference Resources for Translators. Implications for Productivity and Translation Quality." *The Interpreter and Translator Trainer* 6 (2): 219–35. https://doi.org/10.1080/13556509.2012.10798837.

—. 2017. "Investigating Trainee Translators' Views on the Pronunciation of English: a Slovene Perspective," *Linguistica*: Sounds and Melodies Unheard: Essays in Memory of Rastislav Šuštaršič 57 (1): 93–106. https://doi.org/10.4312/linguistica.57.1.93-106.

Hirci, Nataša, Tamara Mikolič Južnič, and Agnes Pisanski Peterlin. (forthcoming). "Enriching Translator Training with Interpreting Tasks: Bringing Sight Translation into the Translation Classroom." *Convergence, Contact and Interaction in Translation and Interpreting Studies*, edited by Eugenia Dal Fovo and Paola Gentile. Berlin: Peter Lang.

Internet World Stats. 2017. Accessed January 22, 2019. https://www.internetworldstats.com/stats7.htm

Jurafsky, Dan, and James H. Martin. 2000. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.* New Jersey: Prentice Hall.

Lambert, Sylvie. 2004. "Shared Attention During Sight Translation, Sight Interpretation and Simultaneous Interpretation." *Meta: journal des traducteurs/Meta: Translators' Journal* 49 (2): 294–306. https://doi.org/10.7202/009352ar.

Li, Xiangdong. 2014. "Sight Translation as a Topic in Interpreting Research: Progress, Problems and Prospects." *Across Languages and Cultures* 15 (1): 67–89. https://doi.org/10.1556/Acr.15.2014.1.4

Mees, Inger M., Barbara Dragsted, Inge G. Hansen, and Arnt Lykke Jakobsen. 2013. "Sound Effects in Translation." *Target* 25 (1): 140–54. https://doi.org/10.1075/target.25.1.11mme.

Mikkelson, Holly. 1994. "Text Analysis Exercises for Sight Translation." In *Vistas: Proceedings of the 31st Annual Conference of ATA*, edited by Peter W. Krawutschke, 381–90. NJ: Learned Information.

Moser-Mercer, Barbara. 1995. "Sight Translation and Human Information Processing." *Basic Issues in Translation Studies. Proceedings of the Fifth International Conference* 2: 159–66.

Moren, Dan. 2018. "Alexa vs. Google Assistant vs. Siri: Google Widens Its Lead." Accessed January 15, 2019. https://www.tomsguide.com/us/alexa-vs-siri-vs-google,review-4772.html

Phillipson, Robert. 2003. *English-only Europe: Challenging Language Policy?* London: Routledge.

Pokorn, K. Nike. 2005. *Challenging the Traditional Axioms: Translation into a Non-Mother Tongue.* Amsterdam/Philadelphia: John Benjamins Publishing.

Pöchhacker, Franz. 2004. *Introducing Interpreting Studies.* London: Routledge.

—. 2010. "The Role of Research in Interpreter Education." *Translation & Interpreting* 2 (1): 1–10.

Rehm, Georg, and Hans Uszkoreit, eds. 2012. *The Slovene Language in the Digital Age.* Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-30636-5.

Riccardi, Alessandra. 2002. "Interpreting Research: Descriptive Aspects and Methodological Proposals." In *Interpreting in the 21st Century: Challenges and Opportunities*, edited by Guiliana Garzone and Maurizio Viezzi, 73–82. Amsterdam/Philadelphia: John Benjamins Publishing.

Sepesy Maučec, Mirjam, Tomaž Rotovnik, Zdravko Kačič, and Janez Brest. 2009. "Using Data-Driven Subword Units in Language Model of Highly Inflective Slovenian Language." *International Journal of Pattern Recognition and Artificial Intelligence* 23 (2): 287–312. https://doi.org/10.1142/S0218001409007119.

Shlesinger, Miriam. 1995. "Shifts in Cohesion in Simultaneous Interpreting." *The Translator* 1 (2): 193–214. https://doi.org/10.1080/13556509.1995.10798957.

Šuštaršič, Rastislav. 2005. *English-Slovene Contrastive Phonetic and Phonemic Analysis and Its Application in Teaching English Phonetics and Phonology.* Ljubljana: Filozofska fakulteta.

Viaggio, Sergio. 1995. "The Praise of Sight Translation (and Squeezing the Last Drop Thereout of)." *The Interpreters' Newsletter* 6: 33–42.

Viezzi, Maurizio. 1990. "Sight Translation, Simultaneous Interpretation and Information Retention." In *Aspects of Applied and Experimental Research on Conference Interpretation*, edited by Laura Gran and Christopher Taylor, 54–60. Udine: Campanotto.

Weber, Wilhelm K. 1990. "The Importance of Sight Translation in an Interpreter Training Program." In *Interpreting: Yesterday, Today, and Tomorrow*, edited by David Bowen and Margareta Bowen, 44–52. Amsterdam: John Benjamins.

Weiner, Sophie. 2016. "Study Says Speech-to-Text Is 3 Times Faster Than Typing On Your Phone." Accessed January 15, 2019. https://www.popularmechanics.com/technology/a22684/phone-dictation-typing-speed/

Young, Holly. 2018. "The Digital Language Divide. How Does the Language You Speak Shape Your Experience of the Internet?" Accessed January 15, 2019. http://labs.theguardian.com/digital-language-divide/

Žgank, Andrej, and Mirjam Sepesy Maučec. 2010. "Razpoznavalnik tekočega govora UMB Broadcast News 2010: nadgradnja akustičnih in jezikovnih modelov." In *Proceedings of the Seventh Language Technologies Conference,* Ljubljana, Slovenia, 28–31. Ljubljana: Institut Jožef Stefan.

Žgank, Andrej, Darinka Verdonik, and Mirjam Sepesy Maučec. 2016. "Razpoznavanje tekočega govora v slovenščini z bazo predavanj SI TEDx-UM." In *Proceedings of the Conference on Language Technologies & Digital Humanities,* Ljubljana, Slovenija, 186–89. Ljubljana: Ljubljana University Press.