Pronunciation practice in Google Translate: focus on French liaison

Kevin Papin[1] and Walcir Cardoso[2]

Abstract. This paper examines the impact of self-directed pronunciation practice using *Google Translate* (GT) on the acquisition of French liaison (the production of latent consonants when they appear in consonant-vowel contexts across words: /z/ in le/za/mis 'the friends'). Second-language (L2) French learners completed homework activities on GT to practice pronunciation, utilizing the tool's Text-To-Speech (TTS) and Automatic Speech Recognition (ASR) capabilities. Using a pretest/posttest design, the acquisition of liaison was assessed in terms of phonological awareness, perception, and production. The results indicated that while the L2 French learners significantly improved in their ability to *produce* liaison, their awareness and perception were not affected due to ceiling effects. Overall, learners reported positive views of their GT-based, self-directed learning experience.

Keywords: Google Translate, automatic speech recognition, text-to-speech synthesis, L2 pronunciation.

1.   Introduction

Speech technologies have been under increasing scrutiny from Computer Assisted Language Learning (CALL) researchers, particularly for their potential to aid pronunciation instruction (Derwing, Munro, & Thomson, 2022). In CALL, TTS synthesizers and automatic ASR have been incorporated into L2 pronunciation teaching to provide machine-generated models to learners, who can then practice their pronunciation by having the software detect their speech and convert it to text (Cardoso, 2018). While empirical research suggests that the use of TTS and

1. Université du Québec à Montréal, Montréal, Canada; papin.kevin@uqam.ca; https://orcid.org/0000-0001-7547-4524
2. Concordia University, Montréal, Canada; walcir.cardoso@concordia.ca; https://orcid.org/0000-0001-6376-185X

ASR can support L2 pronunciation development in English (e.g. Cardoso, 2018; McCrocklin, 2016), only a handful of studies have targeted L2 French (Liakin, Cardoso, & Liakina, 2017). Moreover, two studies have examined the benefits of combining TTS and ASR in a single learning experience (He & Cardoso, 2021; van Lieshout & Cardoso, 2022). GT is one application in which these speech capabilities can be explored for pronunciation instruction.

To examine the use of GT and its built-in speech capabilities for teacher-guided but autonomous pronunciation instruction, we focused on the acquisition of a phonological feature of French: the liaison, or the re-syllabification of latent consonants when they appear in consonant plus vowel contexts (e.g. compare /z/ in liaised 'le/z/ amis' [le.za.mi] 'the friends' with non-liaised 'le/z/ livres' [le.li.vr] 'the books'). This mixed-method study addressed two Research Questions (RQs).

> RQ1: Does the pedagogical use of GT's speech features (TTS, ASR) as part of a homework assignment assist in the learning of French liaison?

> RQ2: What are the L2 learners' views on the proposed GT-based pronunciation practice?

2. Method

Twenty adult L2 French learners (beginners) participated in this case study. Adopting a perspective to L2 research that reflects 'pedagogical realities' (Erlam & Tolosam, 2022), participants were asked to complete a self-directed pronunciation practice using TTS and ASR in GT as part of a one-hour homework activity. The activity targeted French liaison, a phonological phenomenon that poses difficulty to beginner learners, and consisted of three learner-paced components. First, learners were asked to copy-and-paste 13 sentences containing liaisons in GT by pressing the speaker button (TTS) to listen to the synthesized voice (1 in Figure 1). Second, learners were instructed to click on GT's microphone to orally produce the same sentences for ASR conversion to text (2 in Figure 1 below).

To measure the impact of the homework practice on the participants' learning of liaison (RQ1), a pretest was administered before the GT practice, which was followed by immediate and delayed (two weeks later) posttests. These online

tests included multiple-choice questions based on audio recordings, and the learners' audio recordings of target phrases containing the liaison. Participants' phonological development was assessed across three levels of analysis, following Celce-Murcia, Brinton, and Goodwin's (2010) framework for pronunciation instruction: phonological awareness (e.g. *Are [target sentences] pronounced the same? Describe*), perception (e.g. *Listen and check if you hear an example of 'liaison' or 'no liaison'*), and production (e.g. *Read aloud sentences [containing liaison]*). A focus group discussion was also held among four participants to gain insights into their views (RQ2) of the proposed pedagogical experience (e.g. *Do you feel your French pronunciation improved using GT?*).

Figure 1.   Homework activity sample: GT interface
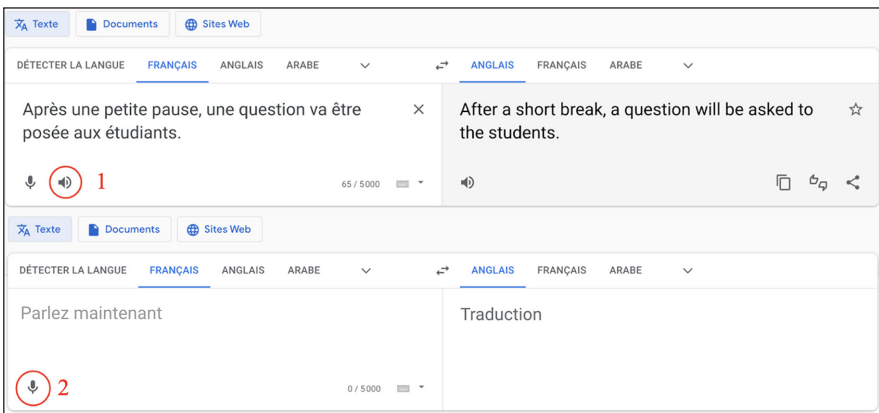
3.   Results

3.1.   Learning gains

No significant changes were observed for awareness, $F(2, 38)=.95$, $p=.621$, partial $\eta2=.03$, or perception, $F(2, 38)=2.26$, $p=.20$, partial $\eta2=.11$. However, the proposed intervention led to statistically significant changes in production over time, $F(2, 38)=11.40$, $p<.001$, partial $\eta2=.38$. Post hoc analysis with a Bonferroni adjustment was conducted, showing a significant increase between the mean scores on the pretest and delayed posttest, 3.04, (95% CI, 1.72 to 4.36), $p<.001$, $d=1.35$. Table 1 reports the evolution of test scores between the pretest and the two posttests, while Figure 2 provides a visual representation of these results, emphasizing the significant values obtained for production.

Table 1.   Test scores: French liaison

	Pretest		Immediate posttest		Delayed posttest	
	M	**SD**	**M**	**SD**	**M**	**SD**
Awareness (/3)	2.55	0.61	2.55	0.51	2.70	0.47
Perception (/16)	9.20	2.35	9.75	1.94	10.40	1.96
Production (/18)	4.64	3.42	5.78	3.73	7.68	3.88

Figure 2.   Production scores (converted to %): French liaison
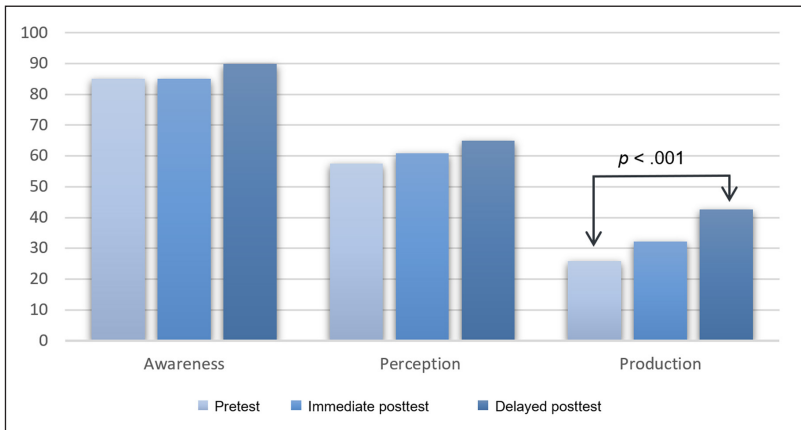
3.2.   Learners' views

Content analysis of the focus group interview indicated that interviewees found their experience 'very useful' overall. Two learners explained that the homework activity 'helped [them] pay more attention' to the liaison and 'raise [their] awareness' to this phonological feature. GT was perceived as a 'robust tool' for L2 pronunciation learning, as the TTS feature can be used to easily check 'how to pronounce' words. Participants felt that the ASR feature has the potential to help them to self-correct, as 'pronunciation must be perfect' to be properly detected by the ASR. Regarding the drawbacks of the experience, two learners found it 'annoying' and 'discouraging' when the ASR could not detect their speech. This sparked comments such as 'I honestly don't plan to [dictate to GT]' and it is 'so more useful for GT to be reading things to me [than me speaking to it]'. After this self-directed pronunciation practice, learners reported they felt encouraged to continue to employ the newly acquired pronunciation learning strategies in their own learning (e.g. to check pronunciation using TTS). However, they did not fully value the pedagogical use of the tool's ASR capability, despite the learning gains described earlier.

4.   Discussion and conclusions

This paper provided initial empirical evidence that the pedagogical use of GT's speech features (TTS and ASR) can support the development of liaison in L2 French (RQ1). While the proposed intervention led to a statistically significant improvement on the production of liaison, it had no impact on awareness and perception. This could be attributed to a ceiling effect in both awareness and perception, as the learners already had advanced knowledge of liaison before the experiment (2.55/3 and 9.2/16 on the pretest respectively, leaving little room for improvement), compared to production (4.64/18 on the pretest). These results are in line with the predictions embedded in Celce-Murcia et al.'s (2010) framework for teaching pronunciation, which posits that L2 phonological features should be taught in the order in which they are acquired, starting with the development of phonological awareness, then progressing toward perception and production. Due to the initial level of liaison knowledge of our participants in awareness and perception, it is not surprising that the proposed treatment only affected their production. Teachers willing to integrate GT to their pronunciation teaching should therefore be mindful of their students' initial phonological knowledge and use GT to target the stage(s) that need(s) improvement.

Due to the naturalistic and self-regulated setting adopted in this study, some external variables might have negatively affected the participants' performance in completing the activities (e.g. internet connection speed, ambient noise at home). The limited number of participants and the fact that this study was based on a one-time intervention should also be acknowledged as a limitation. However, this is also one of the strengths of our study, which constitutes a 'pedagogical reality' (Erlam & Tolosam, 2022) with ecological validity, involving an activity that is common among L2 learners: the completion of out-of-class, teacher-guided, self-regulated homework assignments. Future studies should examine the impact of GT-based pronunciation practice on the development of different phonological features of French (and other languages) and explore learner-GT interactions to examine behavior-effect relations.

5.   Acknowledgments

References

Cardoso, W. (2018). Learning L2 pronunciation with a text-to-speech synthesizer. In P. Taalas, J. Jalkanen, L. Bradley & S. Thouësny (Eds), *Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018* (pp. 165-170). Research-Publishing.net. https://doi.org/10.14705/rpnet.2018.26.806

Celce-Murcia, M., Brinton, D., & Goodwin, J. (2010). *Teaching pronunciation: a course book and reference guide*. Cambridge University.

Derwing, T., Munro, M., & Thomson, R. (2022). (Eds). *The Routledge handbook of second language acquisition and speaking*. Routledge. https://doi.org/10.4324/9781003022497

Erlam, R., & Tolosam, C. (2022). *Pedagogical realities of implementing task-based language teaching*. Benjamins. https://doi.org/10.1075/tblt.14

He, Y., & Cardoso, W. (2021). Can online translators and their speech capabilities help English learners improve their pronunciation? In N. Zoghlamu, C. Brudermann, S. Carré, M. Grosbois, L. Bradley & S. Thouësny (Eds.), *CALL and professionalisation: short papers from EUROCALL 2021* (pp. 127–131). Research-Publishing.net. https://doi.org/10.14705/rpnet.2021.54.1320

Liakin, D., Cardoso, W., & Liakina, N. (2017). The pedagogical use of mobile speech synthesis: focus on French liaison. *Computer Assisted Language Learning*, *30*(3-4), 325-342. https://doi.org/10.1080/09588221.2017.1312463

McCrocklin, S. (2016). Pronunciation learner autonomy: the potential of automatic speech recognition. *System, 57*, 25-42. https://doi.org/10.1016/j.system.2015.12.013.

van Lieshout, C., & Cardoso, W. (2022). Google Translate as a tool for self-directed language learning. *Language Learning & Technology*, *26*(1), 1-19. http://hdl.handle.net/10125/73460

Intelligent CALL, granular systems and learner data: short papers from EUROCALL 2022
Edited by Birna Arnbjörnsdóttir, Branislav Bédi, Linda Bradley, Kolbrún Friðriksdóttir, Hólmfríður Garðarsdóttir, Sylvie Thouësny, and Matthew James Whelpton