EUROCALL 2023: CALL for all Languages
15-18 August 2023, University of Iceland, Reykjavik
DOI: https://doi.org/10.4995/EuroCALL2023.2023.16993

Exploring users' perceptions of ASR for writing narrative texts

Danial Mehdipour-Kolour[a] and Walcir Cardoso[b]

[a]Department of Education, Concordia University, Montreal, danial.mehdipourkolour@mail.concordia.ca and [b]Department of Education, Concordia University, Montreal, walcir.cardoso@concordia.ca

How to cite: Mehdipour-Kolour, D.; Cardoso, W. (2023). Exploring users' perceptions of ASR for writing narrative texts. In *CALL for all Languages - EUROCALL 2023 Short Papers.* 15-18 August 2023, University of Iceland, Reykjavik. https://doi.org/10.4995/EuroCALL2023.2023.16993

Abstract

This study examines how users perceive Automatic Speech Recognition (ASR) as a tool for writing narrative texts, and compares the perceptions of two groups of users: native and non-native English writers. As such, this study aimes to answer the following questions: (1) How do English writers perceive the use of ASR as a writing tool?; and (2) How do native and non-native English writers' perceptions compare in terms of using ASR as a writing tool? To answer these questions, we employed the Technology Acceptance Model 2 (TAM2) to investigate 60 participants' perceptions of utilizing ASR for producing narrative texts. Our findings from analyzing seven components of TAM2 show that writers express a positive attitude towards utilizing ASR as a tool for composing texts. Our findings also indicate no noticeable differences between how native and non-native English writers perceive the usefulness of ASR for creating texts. This is contrary to our hypothesis that native speakers, owing to their more advanced pronunciation skills in English, might have a more favorable attitude towards using ASR.

Keywords: *automatic speech recognition, L2 writing, technology acceptance model, user perception.*

1. Introduction

This study examines the use of ASR as a writing tool and, more specifically, how writiers perceive its use for composing narrative texts, a genre that tells a story in the form of a novel or personal essay. Narrative genre is often considered "cognitively easier and [uses] less complex language than […] exposition and argumentation" (Weigle, 2010, p. 100). ASR identifies voice input and converts it to text output, which is automatically displayed on the screen. Although the utilization of voice as an *input medium* is increasing, particularly in mobile settings (Sohail, 2020), there remains limited knowledge about users' perceptions of using ASR as a writing tool, especially considering new and more advanced speech-based technologies, such as those found in Google Voice Typing.

While research has demonstrated that ASR is useful for enhancing second language (L2) oral skills, such as pronunciation and speaking (Cox & Davies, 2012; Liakin et al., 2015), there is no existing evidence of how adult native and non-native English writers without learning difficulties perceive the use of ASR for writing texts. Moreover, no research has investigated whether there are differences between these two groups of writers in perceiving the usefulness of this technology for writing narrative texts. In fact, excluding a recent study by Johnson and Cardoso (in press) involving only L2 writers, research of this type has not been conducted for approximately two decades (e.g. Leijten & Van Waes, 2005; MacArthur, 1999). In addition, apart from Johnson and Cardoso (in

press), the limited existing studies have examined ASR solely with young learners, e.g. to assist literacy development or writers with learning difficulties (Quinlan, 2004).

Following Cardoso's (2022) chronological framework for conducting CALL research, this study focuses on an early stage of investigation by examining users' perceptions of ASR as a writing tool. As Venkatesh and Davis (2000) indicate, analyzing users' perceptions provides insights into factors influencing acceptance and potential adoption of a target technology. According to the authors, *perceived usefulness* and *ease of use* are key determinants of users' intention to adopt a new technology. To examine users' perceptions, we employed Venkatesh and Davis' (2000) Technology Acceptance Model 2 (TAM2), which is implemented in the form of a survey with a 5-point Likert scale (ranging from 1 = strongly agree to 5 = strongly disagree). The survey included the following seven constructs: (1) Perceived Ease of Use (PEU) (or the extent the user believes that using the technology will be effortless; e.g. *I find ASR easy to use*); (2) Perceived Usefulness (PU) (or the extent the user believes that using the technology will improve their performance; e.g. *Using ASR improves my writing performance*); (3) Intention to Use (IU) (e.g. *I intend to use ASR as a writing tool*); (4) attitude (AT) (e.g. It is a good idea to use ASR for writing purposes); (5) Subjective Norm (SN) (or the perception that is shaped by the influence of external stakeholders like school administrators; e.g. *As a university student, it is important for me to use ASR for writing purposes*); (6) Output Quality (OQ) (or the extent to which the technology can perform the target tasks effectively; e.g. *The quality of ASR-produced text is high*); and (7) Result Demonstrability (RD) (or the extent the results of using the technology are demonstrable; e.g. *I believe I can share my experience using ASR with others*). Each construct of the survey serves as a predictor for the user's perception and acceptance of ASR as a tool for writing.

Given that speech recognition systems, often trained on native speech, may struggle with non-native accents (Liakin et al., 2015), this study moreover compared native and non-native English writers' perceptions of using ASR as a tool for composing narrative texts.

Based on the above, the study was guided by the following Research Questions (RQs):

(1) How do users perceive the adoption of ASR as a tool for producing narrative texts?

(2) How do native and non-native English speakers' perceptions differ regarding the use of ASR as a tool for writing narrative texts?

2. Method

2.1. Context and participants

This study recruited 60 participants, including Native Speakers (NSs; *n*=30) and Non-Native Speakers (NNSs; *n*=30), between the ages of 20 and 50, with a gender distribution of 22 males and 37 females. Participants were undergraduate and graduate students at an Anglophone university in Montreal, Canada. As is customary in research involving humans, consent was obtained from all participants, who were compensated $20 for their participation. They were recruited using social media and other calls for participation, such as email lists. Participants spoke a variety of native languages (e.g. Farsi, French). Due to the COVID-19 pandemic, this study was carried out online using the Zoom video conferencing platform, in individual sessions in which each participant interacted with the ASR for writing their narrative texts. These sessions were video-recorded for analysis (reported in another study).

2.2. Procedures and instruments

Participants were asked to use Google Voice Typing (Google's ASR engine) to produce texts in Google Docs, the application selected for this research. Firstly, participants were assigned free writing activities (e.g. a visual writing prompt) to familiarize themselves with ASR-based writing and to promote fluency in text production (Tynjälä et al., 2001). This phase was followed by a dictation activity, including two short passages. This activity prepared participants for writing longer texts with ASR and helped participants to adjust their writing behavior to punctuate and format their texts, thus becoming more acquainted with ASR features and its writing potential. After practicing

writing with ASR, participants were asked to complete the narrative writing section. Three different narrative prompts were presented (e.g. *success story: write a story about something you accomplished*), and participants were instructed to select two of them and, within 20 minutes, write their narratives using ASR. After the successful completion of the assigned tasks, participants were requested to take part in a 20-minute survey informed by the TAM2 model. The survey was administered online using Google Forms (see sample questions in Section 1), a well-known and reliable platform for performing surveys.

2.3. Data analysis

To find out how users perceive the adoption of ASR as a tool for writing narrative texts (RQ1), a one-sample t-test was administered. Prior to conducting the t-test, a Shapiro-Wilk test was run to assess the normality assumption. The test conducted evaluated whether participants' perception scores were statistically different from *neutral*, which was considered the population perception..

3. Results

Shapiro-Wilk's test revealed that perception scores were normally distributed (p >.05) with no outliers detected. Tables 1 and 2 show the results of the one-sample t-test and descriptive statistics for each unique perception-informed construct examined. It can be asserted that, in general, the participants developed a neutral opinion towards their use of ASR as a writing tool (the lower the Means, the higher the acceptance), even though these results demonstrate the participants' acceptance of several of the TAM2 constructs (i.e. PEU, AT, and RD).

Table 1. Descriptive statistics.

Statistics	**PEU**	**PU**	**IU**	**AT**	**SN**	**OQ**	**RD**
M	2.20	3.08	3.12	2.40	3.87	2.95	1.94
Mdn	2.16	3.00	3.12	2.33	4.00	3.00	1.75
SD	.72	.90	1.05	.89	.88	.84	.97
R	3	3.83	4	4	3	3.67	4

Table 2. One-sample t-test results.

Variable	**t**	**df**	**Sig. (2-tailed)**	**Mean Difference**
PEU	-6.399	59	.000	-.59722
PU	2.479	59	.016	.28889
IU	2.377	59	.021	.32500
AT	-3.453	59	.001	-.40000
SN	9.450	59	.000	1.07500
OQ	1.434	59	.157	.15556
RD	-6.791	59	.000	-.85833

To address the second research question, which explored the differences in perceptions between native and non-native English speakers regarding their use of ASR as a writing tool, a Mann-Whitney U test was conducted. The findings obtained indicate that there was no statistically significant difference between native and non-native English speakers' perceptions across most TAM2 constructs adopted in this study. However, it was observed that native speakers exhibited a higher mean than their counterparts in PEU (see Table 3).

Table 3. Independent-samples Mann-Whitney U test tesults: Comparing native and non-native writers.

Variable	PEU	PU	IU	AT	SN	OQ	RD
Mann-Whitney U	242.00	386.50	321.50	390.50	342.00	395.00	328.50
Wilcoxon W	707.00	851.50	786.50	855.50	807.00	860.00	793.50
Test Statistic	242.00	386.50	321.50	390.50	342.00	395.00	328.50
Standard Error	67.367	67.43	67.36	67.14	66.314	67.13	65.88
*STS	-3.088	-.942	-1.907	-.886	-1.629	-.819	-1.844
Asymptotic Sig.(2-sided test)	.002	.346	.056	.376	.103	.413	.065

4. Discussion

This study examined how users perceive ASR as a tool for writing narrative texts (RQ1) and whether there is a difference in perception among two groups of users: native and non-native English writers (RQ2). Regarding the first question, the results reported indicate that participants' perceptions were slightly more positive compared to average population norms. This suggests that the majority of ASR writers who participated in this research either agreed with or had a neutral stance concerning the utilization of ASR as a tool for composing texts. The presence of such a view can be explained by the user-friendly interface of ASR (van Doremalen et al., 2016). In addition, the use of one's speech fosters fluency (and possibly efficiency) in writing, since it eliminates or minimizes transcription problems, thus alleviating the writer's cognitive burden in the writing process (MacArthur & Cavalier, 2004). However, the results also revealed issues with the ASR tool (observed during the treatment), which may have influenced the participants' sometimes lukewarm responses. Specifically, the adopted ASR system sometimes struggled to accurately recognize certain words when spoken rapidly or inacurately. Additionally, the technology lacked proper text formatting and punctuation without providing verbal cues like "full stop" or "question mark" at the end of each utterance.

For the second research question, our findings showed that native and non-native English writers had comparable views on using ASR as a writing tool. This aligns with recent research indicating that ASR output quality is not substantially impacted by accented speech (see McCrocklin & Edalatishams, 2020), due to recent developments in speech recognition. Therefore, non-native speakers could potentially gain similar benefits from ASR technology as native speakers, due to ASR's ability to accurately transcribe L2-accented speech.

5. Conclusions

The aim of this study was to explore users' perceptions of employing ASR as a tool for writing narrative texts and to find out whether there is a different in perception among native and non-native English writers. The findings suggest ASR engines like Google Voice Typing are beneficial writing aids for both native and non-native English speakers. One possible factor influencing positive user perceptions of ASR-assisted writing is the technology's

ability to accurately transcribe both native and accented speech. These results make important contributions to the literature in two key areas. First, we examined user perspectives on employing ASR for writing, providing valuable insights into this relatively unexplored topic. Second, we focused on comparing native and non-native English speakers, an important yet overlooked subject in prior ASR writing studies. Overall, this work expands current knowledge by eliciting users' views on ASR-facilitated writing while highlighting differences based on pronunciation proficiency of native and non-native speakers of English.

References

Cardoso, W. (2022). Technology for Speaking Development. In T. Derwing, M. Munro, & R. Thomson (Eds), *Routledge Handbook on Second Language Acquisition and Speaking* (p. 299-313). Routledge, Taylor & Francis. https://www.routledge.com/The-Routledge-Handbook-of-Second-Language-Acquisition-and-Speaking/Derwing-Munro Thomson/p/book/9780367900847

Cox, T., & Davies, R. (2012). Using automatic speech recognition technology with elicited oral response testing. *CALICO Journal*, *29*(4), 601–618. https://www.jstor.org/stable/10.2307/calicojournal.29.4.601

Johnson, C. & Cardoso, W. (in press). Hey Google, let's write: Examining L2 learners' acceptance of automatic speech recognition as a writing tool. *CALICO Journal*.

Leijten, M., & Van Waes, L. (2005). Writing with speech recognition: The adaptation process of professional writers with and without dictating experience. *Interacting with Computers*, *17*(6), 736–772. https://doi.org/10.1016/j.intcom.2005.01.005

Liakin, D., Cardoso, W., & Liakina, N. (2015). Learning L2 pronunciation with a mobile speech recognizer: French /y/. *CALICO Journal*, *32*(1), 1–25. https://doi.org/10.1558/cj.v32i1.25962

MacArthur, C. (1999). Overcoming barriers to writing: Computer support for basic writing skills. *Reading and Writing Quarterly*, *15*(2), 169–192. https://doi.org/10.1080/105735699278251

MacArthur, C., & Cavalier, A. (2004). Dictation and speech recognition technology as test accommodations. *Exceptional Children*, *71*(1), 43–58. https://doi.org/10.1177/001440290407100103

McCrocklin, S., & Edalatishams, I. (2020). Revisiting Popular Speech Recognition Software for ESL Speech. *TESOL Quarterly*, *54*(4), 1086–1097. https://doi.org/10.1002/tesq.3006

Perreault, C., & Mathew, S. (2012). Dating the origin of language using phonemic diversity. *PloS One*, *7*(4), 1–8.

Quinlan, T. (2004). Speech recognition technology and students with writing difficulties: Improving fluency. *Journal of Educational Psychology*, *96*(2), 337–346. https://doi.org/10.1037/0022-0663.96.2.337

Selouani, S., Lê, T., Benahmed, Y., & O'Shaughnessy, D. (2008). Speech-enabled tools for augmented interaction in e-learning applications. *International Journal of Distance Education Technologies*, *6*(2), 1–20. https://doi.org/10.4018/jdet.2008040101

Sohail, M. (2020). Exploring user satisfaction with AI-enabled voice-activated smart phone assistants. *PACIS*. https://aisel.aisnet.org/pacis2020

Tynjälä, P., Mason, L., & Lonka, K. (2001). Writing as a learning tool: An introduction. In P. Tynjälä, L. Mason, & K. Lonka (Eds.), *Writing as a learning tool: Integrating theory and practice* 7, 7–22. Springer, Dordrecht. https://doi.org/10.1007/978-94-010-0740-5_2

van Doremalen, J., Boves, L., Colpaert, J., Cucchiarini, C., & Strik, H. (2016). Evaluating automatic speech recognition-based language learning systems: a case study. *Computer Assisted Language Learning*, *29*(4), 833–851. https://doi.org/10.1080/09588221.2016.1167090

Venkatesh, V., & Davis, F. (2000). Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Weigle, S. (2010). Assessing Writing. In *Cambridge University Press*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732997