A comparison of learner characteristics, beliefs, and usage of ASR-CALL systems

Gemma Artieda[1] and Bindi Clements[2]

Abstract. Wall Street English[3] has built online activities that allow students to record phrases and receive word-level Automatic Speech Recognition (ASR) driven pronunciation feedback. Students in language centres in China, Vietnam, Saudi Arabia, and Italy (*N*=2,867) used ASR-Computer Assisted Language Learning (CALL) activities, and some (*N*=482) completed a questionnaire. A high number of students reported that ASR-CALL activities helped them to improve their pronunciation. However, the study found remarkable differences in usage of product features across countries, with students from Vietnam and China using more retries than Saudi Arabia, and students from Italy using the fewest retries. Students from China, Vietnam, and Saudi Arabia more frequently listened to model audios than students from Italy. A series of Kruskal-Wallis tests revealed significant group differences between dominant L1 and students' beliefs and perceptions using ASR, and between age groups and students' beliefs and perceptions using ASR. This study points to the importance of considering regional differences, and suggests that learner engagement may depend not only on the effectiveness of the technology, but also on learner beliefs and perceptions.

Keywords: ASR, speech recognition, pronunciation, pronunciation feedback, learner beliefs and perceptions.

1. Introduction

ASR-CALL activities offer considerable opportunities for individualised practice and personalised feedback on pronunciation (Levis, 2007), and recent studies

1. Wall Street English, Barcelona, Spain; gemma.artieda@wallstreetenglish.com; https://orcid.org/0000-0001-8807-8571
2. Wall Street English, Barcelona, Spain; bindi.clements@wallstreetenglish.com; https://orcid.org/0000-0002-4901-4070
3. Wall Street English is an international school of English for adults.

How to cite this article: Artieda, G., & Clements, B. (2019). A comparison of learner characteristics, beliefs, and usage of ASR-CALL systems. In F. Meunier, J. Van de Vyver, L. Bradley & S. Thouësny (Eds), *CALL and complexity – short papers from EUROCALL 2019* (pp. 19-25). Research-publishing.net. https://doi.org/10.14705/rpnet.2019.38.980

demonstrate that ASR-CALL activities can have a measurable impact on learning (inter alia, Golonka et al., 2014).

Perceptions regarding the difficulty of achieving English pronunciation skills may be linked to L1/nationality (cf. Cenoz & Lecumberri, 1999; Simon & Taverniers, 2011), and while there are commonalities, beliefs regarding effective learning strategies have been found to differ according to learner L1/nationality (Nowacka, 2012). Such beliefs may lead to pre-conceptions relating to the effectiveness of ASR-CALL activities, which may affect student engagement. This study explores adult students' engagement with ASR-CALL activities and aims to address the following research questions.

- Do students think pronunciation activities with ASR help them improve their pronunciation?
- Do students in four countries (China, Vietnam, Italy, and Saudi Arabia) make different use of ASR-CALL activity features?
- Are there differences between L1s and age groups in students' beliefs and perceptions on learning pronunciation using ASR?

2. Method

2.1. Context and participants

Wall Street English uses proprietary learning content to deliver a blended-learning course, offering a combination of multimedia self-study lessons and face-to-face teacher-led classes. It incorporated ASR word-level pronunciation feedback into three multimedia activity types: (1) repeat and practise, (2) read and record, and (3) conversation.

Students were exposed to these ASR activities as part of their course over six weeks. Researchers had access to anonymised back-end data at the end. Participants (*N*=2,867) completed ASR-powered course activities, and 482 of these responded to an optional, anonymous online questionnaire. There was an unequal age range distribution in questionnaire participants in terms of the most common age groups: 19-40 in China, 16-40 in Vietnam, 16-30 in Saudi Arabia, and 23-60 in Italy (Table 1).

Table 1. Participants

	China		Vietnam		Saudi Arabia		Italy		Total
	N	%	N	%	N	%	N	%	N
Studied activities with ASR	1,153	40%	1,192	42%	171	6%	351	12%	2,867
Took the questionnaire	153	32%	173	36%	85	18%	71	15%	482

2.2. Instruments

Learners completed the activities as part of their normal studies. They were familiar with them, only ASR feedback was new. Questionnaire items (see Table 2 and Table 3) were based on previous research. A six-step Likert scale (from 1=not at all to 6=fully agree) was used.

3. Results and discussion

Ninety-five percent of students believed that pronunciation activities with ASR helped improve their pronunciation. Vietnamese students were the most enthusiastic (98.8%), closely followed by students in Italy (98.5%), then in Saudi Arabia (95.2%), and, finally, students in China (91.5%).

Concerning the usage of ASR activity features, students in all territories used a much lower number of attempts than they were allowed. Students in China, Saudi Arabia, and Italy only used one attempt (*Mdn*=1) and those in Vietnam used two (*Mdn*=2), whereas they were allowed three attempts in Activities 1 and 2, and four in 3. For listening to model audio recordings, students in China and Vietnam reported using this feature the most (82%), less so for students in Saudi Arabia (57%) and in Italy (41%). Finally, students in Vietnam reported listening to own utterance recordings the most (80%), less so for students in China (66%) and in Saudi Arabia (61%), and remarkably lower usage was reported by students in Italy (57%).

Finally, we investigated potential between-group differences in age groups and L1s in students' beliefs and perceptions about L2 pronunciation items. The analyses for age revealed statistically significant differences between age groups and all items but one (Item 4 – see Table 2). Differences were favourable to the youngest age group (16-22), suggesting that older learners may have less overall confidence in their ability to acquire pronunciation skills (Marinova-Todd, Marshall, & Snow, 2000).

Table 2. Kruskal-Wallis test results for beliefs and perceptions about L2 pronunciation and age ranges

#	Items	N	M(SD)	Mean Rank	df	χ^2	p	Effect Size d
1	I believe that I will eventually be able to speak English very well.	482			2	49	.000*	0.66
	16-22	167	5.26(.83)	268.45				
	23-30	153	4.97(1.21)	243.62				
	31-60+	162	4.75(1.18)	211.72				
2	I can use technology to help me improve my pronunciation.	482			2	14.4	.001*	0.32
	16-22	167	5.22(.93)	272.34				
	23-30	153	4.73(1.23)	219.60				
	31-60+	162	4.80(1.27)	230.39				
3	I feel at ease when I have to speak English.	482			2	19.3	.000*	0.03
	16-22	167	4.37(1.27)	278.53				
	23-30	153	3.83(1.42)	227.26				
	31-60+	162	3.74(1.39)	216.77				
4	I feel insecure about my pronunciation.	482			2	.9	.624	
	16-22	167	3.67(1.55)	248.46				
	23-30	153	3.60(1.57)	242.02				
	31-60+	162	3.53(1.39)	233.83				
5	It is important for me to speak English with an excellent English pronunciation.	482			2	.21	.000*	0.12
	16-22	167	5.64(.90)	269.93				
	23-30	153	5.43(1.02)	241.87				
	31-60+	162	5.24(1.05)	211.85				
6	I am happy with my pronunciation as long as people can understand me.	482			2	8.4	.015*	0.23
	16-22	167	3.76(1.60)	259.58				
	23-30	153	3.21(1.78)	216.05				
	31-60+	162	3.59(1.55)	246.90				

* $p<0.05$, d_{cohen}=0.2 (small), 0.5 (medium), 0.8 (large)

Results for L1 analyses are displayed in Table 3. The analyses yielded statistically significant differences between L1s and all beliefs and perceptions about L2 pronunciation statements, in line with some findings reported in Nowacka (2012).

The effect sizes were noticeably larger than for age groups, which may help to explain some differences found in feature usage. For example, students from Italy, who reported the lowest instances of listening to both the sample and their own recorded audios, also demonstrated lowest self-belief in ability to be able to speak English very well, and placed lowest importance on being able to speak with excellent pronunciation.

Table 3. Kruskal-Wallis test results for beliefs and perceptions about L2 pronunciation with L1s

#	Items	N	M(SD)	Mean Rank	df	χ^2	p	Effect Size d
1	I believe that I will eventually be able to speak English very well.	482			3	29.5	.000*	0.48
	Chinese	153	4.83(1.37)	234.89				
	Vietnamese	173	5.20(.85)	259.07				
	Arabic	85	5.29(.89)	276.22				
	Italian	71	4.52(.99)	171.37				
2	I can use technology to help me improve my pronunciation.	482			3	19.1	.000*	0.37
	Chinese	153	4.58(1.42)	211.92				
	Vietnamese	173	5.19(.92)	266.75				
	Arabic	85	5.12(1.05)	264.54				
	Italian	71	4.77(1.07)	216.13				
3	I feel at ease when I have to speak English.	482			3	37.9	.000*	0.56
	Chinese	153	3.70(1.52)	217.84				
	Vietnamese	173	4.47(1.21)	289.59				
	Arabic	85	3.92(1.39)	231.39				
	Italian	71	3.49(1.09)	187.42				
4	I feel insecure about my pronunciation.	482			3	13.9	.003*	0.30
	Chinese	153	3.47(1.67)	228.31				
	Vietnamese	173	3.93(1.35)	270.12				
	Arabic	85	3.22(1.66)	208.49				
	Italian	71	3.60(1.35)	239.70				

5	It is important for me to speak English with an excellent English pronunciation.	482			3	113.1	.000*	1.09
	Chinese	153	5.35(1.13)	236.86				
	Vietnamese	173	5.71(.68)	269.73				
	Arabic	85	5.88(.42)	294.76				
	Italian	71	4.45(1.16)	118.96				
6	I am happy with my pronunciation as long as people can understand me.	482			3	77.1	.000*	0.85
	Chinese	153	2.62(1.63)	168.05				
	Vietnamese	173	3.76(1.52)	251.76				
	Arabic	85	4.43(1.51)	315.31				
	Italian	71	4.09(1.19)	286.42				

* $p<0.05$, d_{cohen}=0.2 (small), 0.5 (medium), 0.8 (large)

4. Conclusions

Learners were overwhelmingly positive towards ASR-CALL's potential in helping improve their pronunciation. However, differences in feature usage were observed between students of different L1s/nationalities, which may be related to differences in learner beliefs.

5. Acknowledgements

We would like to thank students and staff in the participating Wall Street English centres.

References

Cenoz, J., & Lecumberri, M. (1999). The acquisition of English pronunciation: learners' views. *International Journal of Applied Linguistics*, *9*(1), 3-15. https://doi.org/10.1111/j.1473-4192.1999.tb00157.x

Golonka, E., Bowles, A., Frank, V., Richardson D., & Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*(1), 70-105. https://doi.org/10.1080/09588221.2012.700315

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics, 27*, 184-202. https://doi.org/10.1017/s0267190508070098

Marinova-Todd, S. H., Marshall, D. B., & Snow, C. E. (2000). Three misconceptions about age and L2 learning. *TESOL Quarterly*, *34*(1), 9-34. https://doi.org/10.2307/3588095

Nowacka, M. (2012). Questionnaire-based pronunciation studies: Italian, Spanish and Polish students' views on their English pronunciation. *Research in Language, 10*(1), 43-61. https://doi.org/10.2478/v10015-011-0048-3

Simon, E., & Taverniers, M. (2011). Advanced EFL learners' beliefs about language learning and teaching: a comparison between grammar, pronunciation, and vocabulary. *English Studies, 92*(8), 896-922. https://doi.org/10.1080/0013838x.2011.604578

Published by Research-publishing.net, a not-for-profit association
Contact: info@research-publishing.net

CALL and complexity – short papers from EUROCALL 2019
Edited by Fanny Meunier, Julie Van de Vyver, Linda Bradley, and Sylvie Thouësny

Publication date: 2019/12/09

Typeset by Research-publishing.net
Cover theme by © 2019 Frédéric Verolleman
Cover layout by © 2019 Raphaël Savina (raphael@savina.net)

Fonts used are licensed under a SIL Open Font License

ISBN13: 978-2-490057-54-2 (Ebook, PDF, colour)
ISBN13: 978-2-490057-55-9 (Ebook, EPUB, colour)
ISBN13: 978-2-490057-53-5 (Paperback - Print on demand, black and white)
Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: décembre 2019.