

YouTube Spoken Corpora in Pronunciation Instruction for Fostering Intercultural Willingness to Communicate

Phi Dat Luu¹

¹ Independent Researcher,
Ho Chi Minh City, Vietnam

✉ ***Corresponding author:**

datphi19030@gmail.com



This work is licensed under CC BY-NC-ND 4.0.
To view a copy of this license, visit
<https://creativecommons.org/licenses/by-nc-nd/4.0>

ABSTRACT: Traditional pronunciation training is largely based on native-speaker ideals, sometimes neglecting the practical realities of worldwide mutual intelligibility. This study investigates how YouTube-based spoken corpora might enhance phonological clarity and encourage EFL students' willingness to communicate across cultures. This study uses a qualitative descriptive synthesis design to examine peer-reviewed empirical literature from the Scopus, Web of Science, Google Scholar and ERIC databases published between 2015 and 2026. The findings indicate that raw YouTube speech contains harmful social prejudices, performative accent stereotypes, and serious demographic mistakes inside automated caption algorithms. Direct digital exposure results in early phonological shock and passive entertainment comfort rather than active, long-term psychological resilience in speaking. The synthesis produces a three-tiered instructional assessment matrix that considers auditory features, visual scaffolding, and tight constraints on cognitive working memory capacity. However, this framework is limited by a heavy reliance on existing literature and Western-centric survey snapshots, necessitating future longitudinal action research and manual human-in-the-loop transcript audits.

KEYWORDS: YouTube spoken corpora, Pronunciation instruction, Mutual intelligibility, Intercultural willingness to communicate, Cognitive load.

How to cite: Luu, P. D. (2026). YouTube spoken corpora in pronunciation instruction for fostering intercultural willingness to communicate. *World Conference on Science, Innovation and Human Progress*. Futurity Research Publishing. <https://doi.org/10.5281/zenodo.21260676>

Introduction

Traditional English pronunciation instruction frequently focuses on the native speaker ideal. However, reciprocal intelligibility is more important than faultless accents in current global communication (Levis, 2020; Wang & Wen, 2023). During real-world encounters, intelligence refers to a listener's ability to comprehend words and understand the speaker's underlying intent. This skill requires both individual sounds and overall speech rhythm (Luu & Nguyen, 2026). Interestingly, learners with better proficiency are more aware of their own faults (Luu & Nguyen, 2026). This self-monitoring shows that pronunciation learning is a kind of intentional mental doing (Luu et al., 2025), though, in practice, building up rhythm, stress, and the real-like flow of speech still feels like a universal struggle for students of a foreign language (Luu & Nguyen, 2026; Petinou & Theodorou, 2019). Language education must consequently transition from inflexible local standards to functional, clear communication in varied circumstances.

Clear pronunciation is not just a linguistic goal. It directly affects a learner's Intercultural Willingness to Communicate (IWTC). IWTC is the psychological readiness to start conversations with people from different cultural backgrounds (Kassing, 1997). In countries like Vietnam, this readiness is a complex negotiation (Luu, 2026). Students often experience a strong tension between wanting to speak and fearing negative social evaluation (Luu, 2026). Anxiety and fear of communication failure create heavy mental barriers (Lu & Hsu, 2008). Many students avoid real-world encounters completely due to their nervousness (Campbell, 2016). True communicative confidence is strongly dependent on a supportive emotional context and good self-regulation. Therefore, language courses must address these emotional hurdles in order to educate students for a global workforce.

Teachers want meaningful feedback in order to bridge the gap between classroom practice and real-world communication. Spoken corpora produced from YouTube are an ideal alternative. These corpora are like huge sets of real video texts and transcripts (Coats, 2023). YouTube material, unlike written textbooks, exposes students to a mix of dialects, natural speech rates, and plenty of cultural symbols (Li & Adnan, 2023; Nguyen & Luu, in press). Platform elements such as comment sections can provide interactive places in which users negotiate meaning. However, there are a few limitations on these digital materials. Automated transcripts tend to be hit by inaccuracies and background noise (Coats, 2023). Besides that, unstructured movies can create cognitive overload, or they might just function as simple amusement. Without explicit instructional supervision, children find it difficult to integrate informal video viewing with formal language improvement.

Research Aim and Research Questions

Despite its benefits, unstructured YouTube content often leads to cognitive fatigue or serves only as amusement (Nguyen & Luu, in press). This problem is exacerbated by gaps in current research: pronunciation studies rely heavily on small samples and subjective self-reports rather than objective acoustic data (Luu & Nguyen, 2026; Petinou & Theodorou, 2019), and literature on IWTC has a strong U.S.-centric bias (Kassing, 1997; Yousaf et al., 2022). A significant gap remains in understanding how to properly filter this vast digital resource to improve voice clarity while reducing cultural discomfort. This qualitative synthesis fills a vacuum by examining how YouTube-based spoken corpora might improve phonological intelligibility and encourage IWTC in EFL learners, offering a clear pedagogical framework for connecting raw digital input to classroom practice. To do this, the analysis is guided by three major questions:

1. How does existing literature evaluate the linguistic features of YouTube-derived spoken data for instructional purposes?
2. How does exposure to authentic, diverse pronunciation features via digital corpora influence learners' intercultural communication apprehension and their willingness to communicate?
3. What pedagogical criteria should be synthesized to develop a selection matrix for YouTube content that balances pronunciation intelligibility and intercultural readiness for EFL learners?

Methodology

This study utilizes an exploratory, qualitative descriptive synthesis design to map how digital media linguistic variables intersect with socio-psychological constructs in language learning (Sandelowski & Barroso, 2007). Moving away from empirical experiments, this approach synthesizes established research patterns to guide university curriculum design (Major & Savin-Baden, 2010). Data collection involved a structured search across Scopus, Web of Science, Google Scholar and ERIC using keywords combining YouTube/spoken corpora, pronunciation attributes (accent diversity, speech rate, intelligibility), and intercultural communication outcomes (IWTC, anxiety). The final dataset was filtered to include peer-reviewed empirical studies (2008–2026) in higher-education EFL/EAP contexts. Non-authentic digital media, dissertations, and K-12 educational studies were systematically excluded.

The selected articles were evaluated through qualitative content analysis and thematic synthesis to build a unified pedagogical framework (Schreier, 2012; Thomas & Harden, 2008). First, initial coding extracted findings regarding the linguistic value of YouTube speech rates and accent diversity. Second, axial coding mapped how exposure to these features alters learners' intercultural communication apprehension under MacIntyre's (1998) WTC model. Finally, these insights were translated into a three-tiered Pedagogical Evaluation Matrix (phonological properties, affective barriers, and intercultural readiness) for university materials development. Rigor and trustworthiness were maintained via a detailed database audit trail and peer debriefing to reach 100% coding consensus (Lincoln & Guba, 1985).

Research Results

Evaluation of linguistic features in YouTube-derived spoken data for instructional purposes

YouTube videos offer real-world accents that challenge sterile textbook English (Watkins & Wilkins, 2011). However, the literature shows these spaces are not neutral. Instead, they act as digital arenas where viewers openly judge and mock non-standard speech in the comments (Cutler, 2020). This creates a classroom trap. Popular videos like accent challenges rely on fake, exaggerated performances rather than real communication (Rymes & Leone-Pizzighella, 2018). They focus on shallow stereotypes that link regional dialects to low social status. This reinforces native-speaker bias instead of building mutual understanding. Without clear teacher guidance, raw YouTube data can harm student attitudes by normalizing social hierarchies under the guise of authentic speech.

Teachers often rely on YouTube's automatic captions, but this creates a false sense of accuracy. These algorithms contain heavy demographic biases. Because standard AI models are trained on narrow data, they make far more errors on female speakers and regional dialects (Tatman, 2017). This technical flaw means raw

captions regularly erase or misinterpret the exact features learners need to study, like code-switching or thick accents (Maxwell-Smith & Foley, 2023). Treating unedited captions as perfect mirrors of speech clarity exposes students to broken text models. It forces diverse speech into rigid, standardized molds. The technology meant to capture reality actually distorts it, so human editing is vital to keep automated bias out of the classroom.

YouTube successfully captures how real people vary their speech speed across different regions and video genres (Coats, 2020). However, this real-world speed creates a direct conflict with learner brain limits. Teachers can use timestamp tools to filter videos by word count (Lin, 2022), but this does not stop the mental exhaustion that fast, continuous speech causes (Juma, 2021). Old textbooks overprotect students with slow audio, leaving them unready for real life. Yet, throwing raw YouTube speeds into a lesson simply replaces safety with total confusion. Fast speech alone does not build skill. Without a strict filtering system based on student levels, raw digital speech overloads short-term memory and paralyzes comprehension.

Exposure to authentic diverse pronunciation features via digital corpora and its influence on intercultural communication apprehension and willingness to communicate

The literature views raw digital corpora as a solution for textbook artificiality, yet this exposure results in an instant psychological dilemma. While different pronunciation traits disrupt the native-speaker monopoly (Levis, 2020), they also increase rather than decrease intercultural fear. Global accents are not freeing for learners; rather, they cause cognitive disorder (Coats, 2020). When pupils are confronted with raw auditory evidence, the idealized expectation of excellent English causes structural shock. YouTube's multimodal features may provide a visual safety net (Benson, 2015), but the comfort is illusory. It covers a more serious anxiety: the worry of social marginalization as a result of real-time decoding failure. As a result, exposure alone does not alleviate communication anxiety; rather, it reveals the frail underpinnings of classroom-trained comprehension.

Current research assumes a smooth, linear progression from phonological shock to emotional ease, but this desensitization is primarily a pedagogical illusion. Listening to successful non-native speakers on digital platforms is thought to increase a learner's self-efficacy (Zadorozhnyy & Lee, 2025). However, this psychological transformation is completely passive and situational. Watching a video does not imply admitting one's own pronunciation mistakes; instead, it normalizes the flaws of those on screen. When the learner moves away from the spectator position, the positive emotions praised in informal learning research frequently fade (Lee et al., 2024). As a result, the literature mixes entertainment-driven comfort with true psychological resilience, creating an unfilled gap between passive acoustic tolerance and active speaking confidence.

Lowering phonological anxiety by digital exposure acts as a direct stimulus for IWTC. By frequently processing real pronunciation qualities, learners transform their ideal L2 self into a skilled global communicator rather of a phony native speaker (Lee & Chiu, 2023). This psychological change increases their willingness to begin discussions in public digital areas such as video comment sections or online language threads (Lee & Hsieh, 2019). However, a key structural tension remains unsolved in contemporary study. Most research use cross-sectional survey data, which only captures a glimpse of learner confidence (Zadorozhnyy & Lee, 2025). Current research does not demonstrate that a student who is mentally prepared to understand varied accents online would keep that same willingness during a real-time, face-to-face intercultural crisis (Luu, 2026).

Pedagogical criteria for developing a selection matrix of YouTube content balancing pronunciation intelligibility and intercultural readiness for EFL learners

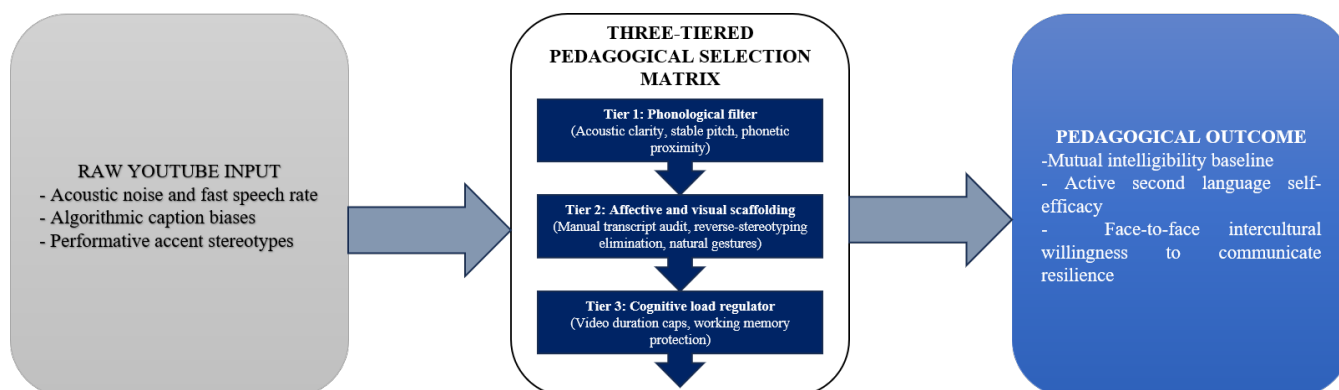
Selection matrices must reject native-speaker models as the default standard. Real communication requires mutual intelligibility rather than native likeness (Chan, 2021). The matrix must prioritize physical properties like clear vowel duration and stable pitch frequencies (Kang et al., 2025; Liu et al., 2014). High background noise prevents learners from adjusting to new acoustic features (Clopper & Bradlow, 2008). Therefore, the initial filter must screen out poor audio quality and severe distortions. The primary goal is a clean speech signal that human ears can decode. Random exposure to diverse accents creates cognitive chaos instead of intercultural readiness. The selection matrix must track accent varieties based on the learner's linguistic background (Olatoye & Olurinola, 2026). Students decode accents better when they share a phonetic space with the speaker (Fishero et al., 2023). The matrix should introduce common sound variations before presenting unique, highly deviant phonemes (Fishero et al., 2023). Unregulated data dumps limit perceptual growth (Mora & Fouz-González, 2024). Instructors must systematically balance familiar and unfamiliar speech features. Accent diversity requires strong visual and textual support to be pedagogically useful (Sardegna & McGregor, 2013).

However, automated captions contain heavy biases against female voices and regional dialects (Tatman, 2017). Reading broken captions while listening also creates a severe attentional split (Wisniewska & Mora, 2017). Tracking phonetic forms and text simultaneously triggers rapid cognitive overload (Wisniewska & Mora, 2020). The matrix must select videos where natural gestures and clear visual contexts ease the processing burden. Intercultural training requires the active removal of linguistic prejudice. Social expectations and ethnic stereotypes often override objective speech clarity (Babel & Mellesmoen, 2019). Through reverse linguistic stereotyping, listeners label clear speech as unintelligible based solely on the speaker's perceived identity (Rubin, 2011). YouTube often hosts performative, mocked accent videos that reinforce social hierarchies. The matrix must include a strict filter to reject content that exploits dialects for entertainment. It must prioritize diverse speakers in professional or academic roles.

The final matrix criterion must measure the mental effort required to process raw digital speech. Accented speech increases cognitive load and slows down acoustic processing times (Mattys & Wiget, 2011). Non-native listeners lose their processing advantage when a speaker's first language is unfamiliar (Liu & Reed, 2018). Simple intelligibility does not guarantee comprehension during complex academic tasks (Jensen & Thøgersen, 2017). The matrix must strictly cap video duration and task difficulty based on human working memory limits (Peng et al., 2020).

Figure 1

The proposed multi-tiered framework for integrating YouTube spoken corpora in EFL pronunciation pedagogy



Conclusions

This qualitative synthesis shows that raw YouTube exposure does not automatically build pronunciation competence or intercultural readiness. Instead, unguided digital input triggers phonological shock, cognitive fatigue, and exposure to platform-driven accent biases. True progress requires shifting from native-speaker ideals to a structured, human-edited selection matrix that strictly balances acoustic clarity, visual scaffolding, and cognitive load limits. However, this framework is currently limited by a lack of direct classroom experiments, a strong Western bias in existing communication literature, and a reliance on brief survey snapshots. To fix these gaps, future university research must utilize longitudinal action research to track real behavioral changes. Educators must also perform manual video transcripts audits to eliminate algorithmic errors and deploy multi-tiered evaluation matrices that cap video lengths to protect human working memory.

References

- Babel, M., & Mellesmoen, G. (2019). Perceptual adaptation to stereotyped accents in audiovisual speech. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 1004-1048). Melbourne, Australia. https://assta.org/proceedings/ICPhS2019Microsite/pdf/full-paper_815.pdf
- Benson, P. (2015). *Commenting to learn: Evidence of language and intercultural learning in comments on YouTube videos*. *Language Learning & Technology*, 19(3), 88-105. <https://scholarspace.manoa.hawaii.edu/bitstreams/3e9342f5-3042-4834-a37b-34ae4518c974/download>
- Campbell, N. (2016). Ethnocentrism and intercultural willingness to communicate-A study of New Zealand management students. *Journal of Intercultural Communication*, 16(1), 1-15. <https://doi.org/10.36923/jicc.v16i1.707>
- Chan, V. (2021). Factors influencing intelligibility and comprehensibility: a critical review of research on second language English speakers. *Journal of English Learner Education*, 12(1), 1-30. <https://stars.library.ucf.edu/jele/vol12/iss1/6/>
- Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech*, 51(3), 175-198. <https://doi.org/10.1177/0023830908098539>
- Coats, S. (2020). Articulation rate in American English in a corpus of YouTube videos. *Language and Speech*, 63(4), 799-831. <https://doi.org/10.1177/0023830919894720>
- Coats, S. (2023). Dialect corpora from YouTube. *Language and linguistics in a complex world*, 79-102. <https://doi.org/10.1515/9783111017433-005>
- Cutler, C. (2020). Metapragmatic comments and orthographic performances of a New York accent on YouTube. *World Englishes*, 39(1), 36-53. <https://doi.org/10.1111/weng.12444>
- Fishero, S., Sereno, J. A., & Jongman, A. (2023). Perception and production of Mandarin-accented English: The effect of degree of accentedness on the Interlanguage Speech Intelligibility Benefit for Listeners (ISIB-L) and Talkers (ISIB-T). *Journal of Phonetics*, 99, 1-18. <https://doi.org/10.1016/j.wocn.2023.101255>
- Jensen, C., & Thøgersen, J. (2017). Foreign accent, cognitive load and intelligibility of EMI lectures. *Nordic Journal*

- of English Studies*, 16(3), 107-137. <https://doi.org/10.35360/njes.414>
- Juma, M. J. (2021). Developing English pronunciation through animation and YouTube videos. *Arab World English Journal*, 12(4), 401-414. <https://doi.org/10.24093/awej/vol12no4.26>
- Kang, O., Bae, Y., & Miao, Y. (2025). L2 speech intelligibility. *Language Teaching*, 1-18. <https://doi.org/10.1017/s0261444825101006>
- Kassing, J. W. (1997). Development of the intercultural willingness to communicate scale. *Communication Research Reports*, 14(4), 399-407. <https://doi.org/10.1080/08824099709388683>
- Lee, J. S., & Chiu, M. M. (2023). Modeling EFL learners' willingness to communicate: The roles of face-to-face and digital L2 communication anxiety. *Annual Review of Applied Linguistics*, 43, 64-87. <https://doi.org/10.1017/s0267190523000090>
- Lee, J. S., Yeung, N. M., & Osburn, M. B. (2024). Foreign language enjoyment as a mediator between informal digital learning of English and willingness to communicate: A sample of Hong Kong EFL secondary students. *Journal of Multilingual and Multicultural Development*, 45(9), 3613-3631. <https://doi.org/10.1080/01434632.2022.2112587>
- Levis, J. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, 6(3), 310-328. <https://doi.org/10.1075/jslp.20050.lev>
- Li, J., & Adnan, H. M. (2023). Exploring the role of productive audiences in promoting intercultural understanding: A study of Chinese YouTube influencer Liziqi's channel. *Studies in Media and Communication*, 11(4), 173-184. <https://doi.org/10.11114/smc.v11i4.6021>
- Lin, P. (2022). Developing an intelligent tool for computer-assisted formulaic language learning from YouTube videos. *ReCALL*, 34(2), 185-200. <https://doi.org/10.1017/s0958344021000252>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.
- Liu, C., Jin, S. H., & Chen, C. T. (2014). Durations of American English vowels by native and non-native speakers: Acoustic analyses and perceptual effects. *Language and Speech*, 57(2), 238-253. <https://doi.org/10.1177/0023830913507692>
- Liu, D., & Reed, M. (2018). Asymmetrical cognitive load imposed by processing native and non-native speech. *Pronunciation in Second Language Learning and Teaching Proceedings*, 10(1), 1-12. <https://www.iastatedigitalpress.com/psllt/article/id/15375/>
- Lu, Y., & Hsu, C. F. (2008). Willingness to communicate in intercultural interactions between Chinese and Americans. *Journal of Intercultural Communication Research*, 37(2), 75-88. <https://doi.org/10.1080/17475750802533356>
- Luu, P. D. (2026). Challenges and strategies of intercultural willingness to communicate among English majors at a Vietnamese public university. *TNU Journal of Science and Technology*, 231(04), 422-429. <https://doi.org/10.34238/tnu-jst.13650>
- Luu, P. D., & Nguyen, P. T. (2026). Exploring common pronunciation errors faced by English majors at a Vietnamese public university. *Journal of Foreign Language Teaching and Learning*, 11(1), 194-228.

<https://doi.org/10.18196/ftl.v11i1.29346>

- Luu, P. D., Anh, N. H. M., & Huong, N. G. (2025). Internal and external determinants of English majors' attitudes towards pronunciation learning at a Vietnamese public university. *Thu Dau Mot University Journal of Science*, 7(4), 813-833. <https://doi.org/10.37550/tdmu.EJS/2025.04.679>
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545-562. <https://doi.org/10.1111/j.1540-4781.1998.tb05543.x>
- Major, C.H., & Savin-Baden, M. (2010). *An Introduction to Qualitative Research Synthesis: Managing the Information Explosion in Social Science Research* (1st ed.). Routledge. <https://doi.org/10.4324/9780203497555>
- Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, 65(2), 145-160. <https://doi.org/10.1016/j.jml.2011.04.004>
- Maxwell-Smith, Z., & Foley, B. (2023). Automated speech recognition of Indonesian-English language lessons on YouTube using transfer learning. In O. Serikov, E. Voloshina, A. Postnikova, E. Klyachko, E. Vylomova, T. Shavrina, E. Le Ferrand, V. Malykh, F. Tyers, T. Arkhangelskiy, & V. Mikhailov (Eds.), *Proceedings of the Second Workshop on NLP Applications to Field Linguistics* (pp. 1-16). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.fieldmatters-1.1>
- Mora, J. C., & Fouz-González, J. (2024). Contrastive input enhancement in captioned video for L2 pronunciation learning. *Audiovisual Input and Second Language Learning*, 61, 150-175. <https://doi.org/10.1075/llt.61.07mor>
- Nguyen, T. T. T., & Luu, D. P. (in press). A study on the perceived effectiveness of YouTube in listening skills among English majors at a Vietnamese Public University. *HO CHI MINH CITY OPEN UNIVERSITY JOURNAL OF SCIENCE-SOCIAL SCIENCES*. <https://journalofscience.ou.edu.vn/index.php/soci-en/article/view/4654>
- Olatoye, O., & Olurinola, O. (2026). How effective are computer-assisted pronunciation and Youtube-based instructions in improving oral english proficiency among secondary school students?. *Language, Technology, and Social Media*, 4(1), 1-18. <https://doi.org/10.70211/ltsm.3026-7196.73>
- Peng, X., Chen, H., Wang, L., Tian, F., & Wang, H. (2020). Talking head-based L2 pronunciation training: Impact on achievement emotions, cognitive load, and their relationships with learning performance. *International Journal of Human-Computer Interaction*, 36(16), 1487-1502. <https://doi.org/10.1080/10447318.2020.1752476>
- Petinou, K., & Theodorou, E. (2019). Promoting speech intelligibility through phonologically dense targets. *Clinical Linguistics & Phonetics*, 33(10-11), 978-990. <https://doi.org/10.1080/02699206.2019.1597168>
- Rubin, D. (2011). The power of prejudice in accent perception: Reverse linguistic stereotyping and its impact on listener judgments and decisions. *Pronunciation in Second Language Learning and Teaching Proceedings*, 3(1), 1-7. <https://www.iastatedigitalpress.com/psllt/article/id/15173/>

- Rymes, B., & Leone-Pizzighella, A. (2018). YouTube-based accent challenge narratives: Web 2.0 as a context for studying the social value of accent. *International Journal of the Sociology of Language*, 2018(250), 137-163. <https://doi.org/10.1515/ijsl-2017-0058>
- Sandelowski, M., & Barroso, J. (2007). *Handbook for synthesizing qualitative research*. Springer Publishing Company.
- Sardegna, V. G., & McGregor, A. (2013). Scaffolding students' self-regulated efforts for effective pronunciation practice. *Pronunciation in Second Language Learning and Teaching Proceedings*, 4(1), 1-12. <https://doi.org/10.1002/9781394259663.ch6>
- Schreier, M. (2012). *Qualitative content analysis in practice*. SAGE Publications.
- Tatman, R. (2017). Gender and dialect bias in YouTube's automatic captions. In D. Hovy, S. Spruit, M. Mitchell, E. M. Bender, M. Strube, & H. Wallach (Eds.), *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing* (pp. 53-59). Association for Computational Linguistics. <https://doi.org/10.18653/v1/W17-1606>
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(1), 1-10. <https://doi.org/10.1186/1471-2288-8-45>
- Wang, Y., & Wen, X. (2023). Nativeness versus intelligibility as goal of English pronunciation teaching in China: Changing attitudes in national syllabi and curriculum standards. *Asian-Pacific Journal of Second and Foreign Language Education*, 8(1), 1-26. <https://doi.org/10.1186/s40862-023-00189-2>
- Watkins, J., & Wilkins, M. (2011). Using YouTube in the EFL classroom. *Language Education in Asia*, 2(1), 113-119. https://doi.org/10.5746/leia/11/v2/i1/a09/watkins_wilkins
- Wisniewska, N., & Mora, J. C. (2017). Pronunciation learning through captioned videos. *Pronunciation in Second Language Learning and Teaching Proceedings*, 9(1), 1-12. <https://www.iastatedigitalpress.com/psllt/article/id/15355/>
- Wisniewska, N., & Mora, J. C. (2020). Can captioned video benefit second language pronunciation?. *Studies in Second Language Acquisition*, 42(3), 599-624. <https://doi.org/10.1017/S0272263120000029>
- Yousaf, M., Ahmad, M., Ji, D., Huang, D., & Raza, S. H. (2022). A cross-cultural comparison of ethnocentrism and the intercultural willingness to communicate between two collectivistic cultures. *Scientific Reports*, 12(1), 1-12. <https://doi.org/10.1038/s41598-022-21179-3>
- Zadorozhnyy, A., & Lee, J. S. (2025). Informal digital learning of English and willingness to communicate in a second language: Self-efficacy beliefs as a mediator. *Computer Assisted Language Learning*, 38(4), 669-689. <https://doi.org/10.1080/09588221.2023.2215279>