Integrating Natural Language Processing and Multimedia Databases in CALL Software: Development and Evaluation of an ICALL Application for EFL listening Comprehension

Danh Cong Vu (martinkongzanh@gmail.com)
Suranaree University of Technology, Thailand

Andrew-Peter Lian (andrew.lian@andrewlian.com)
Suranaree University of Technology, Thailand

Peerasak Siriyothin (peerasak@sut.ac.th)
Suranaree University of Technology, Thailand

Abstract

This paper presents an original computer-assisted language learning (CALL) app for EFL listening comprehension. The software, *Listening Hacked*, utilized a multimedia database and natural language processing (NLP) to create a personalized, autonomous learning environment for EFL learners. The paper is organized into two parts. In the first part, the paper describes the development of the software, including its theoretical underpinnings and developed functionalities. The second part, reports on the evaluation of the software which involved an experiment with 53 English-major Vietnamese students. The students were randomly assigned to the Experimental Group (EG) and Control Group (CG). The EG learned EFL listening by watching English-speaking movies, doing paused transcription tasks, and using various help options available on the platform to complete the tasks. The CG learned by doing traditional listening exercises with comprehension questions on Google classroom. The t-test and repeated measures of ANOVA results indicate that after 12 weeks of study, the students in the EG showed improvement in their EFL listening performance and that they performed better on the posttest than those who learned with traditional methods. The paper also discusses some implications of the findings in the context of researching, developing, and implementing CALL software for L2 listening.

Keywords: L2 listening comprehension; ICALL; personalized learning; multimedia; natural language processing

Introduction

Listening plays an essential part in human communication and contributes to the development of other communication skills. Research in communication has unveiled that listening can take up to 55% of our daily communication activities (Worthington & Fitch-Hauser, 2018). In the field of second/foreign language (L2) education, listening has been corroboratively considered as a primary source for L2 acquisition (Rost, 2001;

Weaver, 1972; Wolvin et al., 2000), dating back to Krashen's input hypothesis (Krashen, 1985) in the 1980s. Despite its fundamental role in language learning, L2 listening is a challenging skill to teach and learn. According to Field (2008), the challenges of L2 listening could be attributed to the complex nature of the skill itself (Wolvin, 2009), the unproductive tradition of instruction, and the lack of interdisciplinary applications. Another important issue that further complicates L2 listening instruction is learner differences. Various factors may influence how individuals listen and comprehend, including but not limited to, gender (Phillips et al., 2001), age (Wolvin, 2018), linguistic knowledge (Long, 1990), and working memory (H. Sakai, 2018). These individual variables pose considerable challenges to the success of any instructional approach to L2 listening, and they inevitably upset the one-size-fits-all approaches often used. An effective instructional approach for L2 listening, therefore, should involve personalization, i.e. tailoring instructional support for individual learning needs and addressing learners' listening difficulties on an individual basis.

L2 practitioners often seek ways of personalizing learning from computer-assisted language learning (CALL) applications. CALL applications can contribute to the effort of personalizing learning in two ways: providing access to substantial learning materials for individual learners, and more importantly, giving individualized support through help options, generally defined as "embedded application resources that assist learners in performing computing operations and/or support language learning" (Cárdenas-Claros & Gruba, 2009, p. 69). Typical help options in CALL environments include dictionaries, transcripts, captions, media replay, speed control, and explanatory feedback (Pujolà, 2002). Research into help options in CALL software for L2 listening has shown that learners express preferences for learning software with help options and that, with use, their comprehension is aided (Cárdenas-Claros & Gruba, 2014; Grgurović & Hegelheimer, 2007; Rivens Mompean & Guichon, 2009).

Despite the abovementioned strengths, existing CALL applications for L2 listening still have some limitations, mostly related to the lack of interdisciplinary applications. Firstly, while advances in technology and infrastructure have fostered access to quality sources of multimedia, most traditional CALL applications providing audio-only materials are still used extensively in many L2 classrooms in Vietnam. Research has provided strong support for multimedia-based CALL software for L2 listening (Brett, 1995; Chi et al., 2022; Diao et al., 2007; Fan, 2020; Meskill, 1996; Pangaribuan et al., 2017; Wahyuni & Septianasari, 2020), notably the use of videos in replacement of audio texts (Lesnov, 2022; Mathew & Alidmat, 2013; Sarani et al., 2014). One of the most prominent theories promoting the use of multimedia in learning is Mayer's (2009) Multimedia Learning theory which states that "people learn better from words and pictures than from words alone" (p.4). Besides, using multimedia materials in L2 listening instruction can facilitate the comprehension process (Al-Athwary & Lasloum, 2021; Guichon & McLornan, 2008; Liang, 2020; Shalmani, 2020; Sueyoshi & Hardison, 2005; Wagner, 2010), and motivate the learners (Cross, 2018).

Secondly, natural language processing (NLP), which works with natural language understanding and natural language generation (Tschichold & Schulze, 2016), while having been already applied in various L2 fields to produce intelligent CALL (ICALL) applications (Azizinezhad & Hashemi, 2013; Heift, 2012), has rarely been found in CALL applications for L2 listening (Dizon, 2020). Despite its great potential, the contributions of ICALL software to L2 listening have remained largely unknown due to either the scarce availability of ICALL software particularly designed for L2 listening and the lack of research on ICALL for L2 listening (Cardenas-Claros & Gruba, 2013).

This paper describes an original ICALL application, *Listening Hacked*, which integrates NLP and a multimedia database to develop a personalized learning environment for EFL listening comprehension. The paper will first describe the app, including the principles underlying its development and the learning activities on the app. The paper will then present empirical evidence relating to its effectiveness in developing the students' EFL listening ability. Specifically, the following question was addressed: What was the effectiveness of Listening Hacked on the students' EFL listening comprehension?

Literature Review

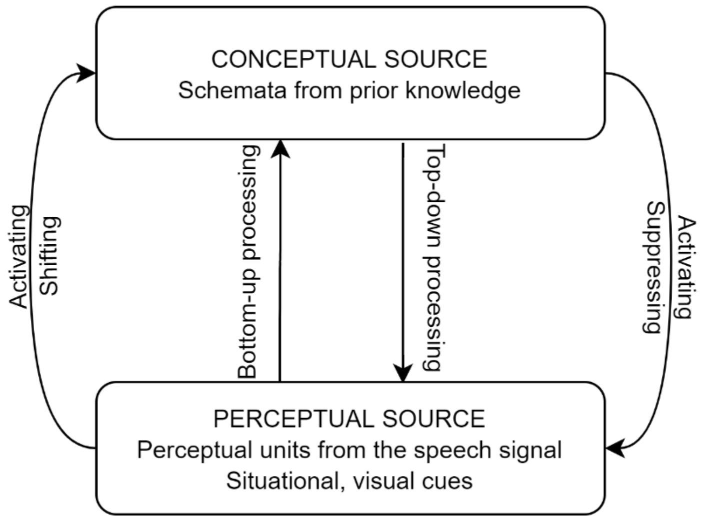
Listening comprehension process

Speech contains continuous streams of acoustic signals which are sent to the listener's ears for perception. We are constantly receiving sound signals around us but not all of them are equally attended to and processed (Worthington & Bodie, 2017). Even when they are attended to, a continuous stream of speech signals is not normally processed as a whole, but rather in smaller units. These units of perception can be features, phonemes, syllables, words, chunks of words, or other sound patterns. Furthermore, these units can be related to each other when they co-occur frequently in certain situations. For example, the words *school, student, teacher, blackboard* can be associated with one another because they co-occur frequently in the contexts of schooling or education. These associations are stored in the listener's mental structures called schemata (Rost, 2006; Goh & Vandergrift, 2021).

According to the Parallel Distributed Processing model (Brodbeck et al., 2022; Scovel, 1998), when we listen, the speech signal can activate potentially corresponding perceptual units at different levels simultaneously (e.g. phoneme level or/and word level) and the schemata relevant to the activated perceptual units. Alternatively, schemata can be activated by non-verbal cues such as situational contexts or visual information (Nguyen & Newton, 2018). In the comprehension process, this activation of schemata, either provoked by verbal or nonverbal cues, serves as a structure of meaning representation onto which activated perceptual information is mapped, according to the structure-building model (Gernsbacher, 1990; Lennox et al., 2020). The 'goodness of fit' will suppress all activated units except for only one. The 'best fit' perceptual information will then be enriched by further evidence from later processing (Rost, 2015; Yang et al., 2022). The structure-building model of comprehension, while originating in reading comprehension research, claims to be cognitive general and is not tied to a specific modality (Gernsbacher, 1990), and thus offers an explanation for individual differences in comprehension skill (McNamara & Magliano, 2009). The distinction between a competent and a less-skilled listener lies in their efficiency of suppressing 'unfit' activated units of perception to arrive at the 'best fit' quickly, instead of their capability in obtaining and enriching contextual information quickly (Gernsbacher & Faust, 1991).

In this study, listening comprehension is conceptualized as a process of constructing coherent mental structures of an attended speech, by activating relevant perceptual units which are then enriched or suppressed in relation to their goodness of fit for the structures (Gernsbacher, 1990, 1991, 1997). The perceptual information is primarily derived from the speech signal in the speaker's utterances while the conceptual information is provided by the activation of relevant schemata. The interaction between the two sources of information, usually mentioned in the literature as the interactive model of listening (Field, 1999; Flowerdew & Miller, 2005; Goh, 2016), generates evidence provided for building the structure of meaning of the attended speech (see Figure 1). Failure to obtain information from one source may lead to failure in comprehension. It is important to stress that while perceptual information derives from the speech signal, the resulting perception is still a product of a combined processing of both perceptual and conceptual information.

Figure 1
Listening comprehension process



Personalized learning

Personalization in education is not a new thing, and yet, it has become a trend in the 21st century. The original goal of personalized education concerns tailoring instruction and providing learning conditions meeting individual learners' needs. This idea leads to the concept of precision education (Cook et al., 2018; Luan & Tsai, 2021; Makhluf, 2020) and more recently precision language education proposed by Lian and Sangarun (2017). The concept of precision language education, inspired by precision education and formerly precision medicine, is a new conceptual move in which language instructions or pedagogical interventions should focus on addressing the actual problems and learning needs experienced by each learner when trying to solve specific language tasks (Bhutoria, 2022; Cook et al., 2018) and moving away from group characteristics, just as precision medicine is intended for specific needs of individual patients.

The move toward precision language education turns individual differences, usually considered as an unavoidable nuisance in education, to that of navigating effective ways to language learning by responding to individual variability instead of normalizing it. Lian and Sangarun (2017) suggest that the distinction between precision education and personalized education is in the goal of getting information “as detailed and accurate as possible” (p.3) about learner variables and needs, with the former being full personalization and “the ultimate objective of the research effort” (p.6).

In precision language education principles, pedagogical intervention should aim at “providing accurate, detailed, timely, adaptive and contextualised personalised data” (Lian & Sangarun, 2017, p. 4) so as to arrive at the most effective solutions to individual learners’ problems. The premise is that we can increase chances of success in language learning if ‘precise’ data of performances can be obtained and made available to learners since awareness of the viability of knowledge construction can trigger effective learning (von Glasersfeld & Cobb, 1983).

Learner Autonomy and L2 learning

As discussed above, successful personalization in learning requires obtaining precise learner performance data. However, in normal teaching situations, the achievement of ‘precision’ might be a challenge unless there is assistance from available technologies. Technology has the potential to collect learner performance data and provide personalized solutions for individual learners based on the collected data. However, whether those technology-based, personalized solutions are effective in assisting learners with their problems still depends on learner autonomy (Reinders, 2018).

Learner autonomy is defined as the learner’s “capacity to take control of one’s own learning” (Benson, 2013, p. 58). As Palfreyman and Benson (2018) pointed out, this learner characteristic is both behavioral and psychological. The psychological aspect of autonomy is evident in the learner motivation and efforts in their learning, while the learner control over their learning activities and cognitive processes are considered behavioral.

One important issue concerning research on autonomy is whether developing autonomy could lead to better L2 learning. While research has not been conclusive about this, several findings (Çetinkaya & Tilfarlioglu, 2020; Dafei, 2007; Daflizar et al., 2022; Hashemian & Soureshjani, 2011; Ngurah & Myartawan, 2013; S. Sakai & Takagi, 2009) including a longitudinal study by Little et al. (2017) support the positive relationship between learner autonomy and L2 learning.

The use of AI in L2 listening pedagogy

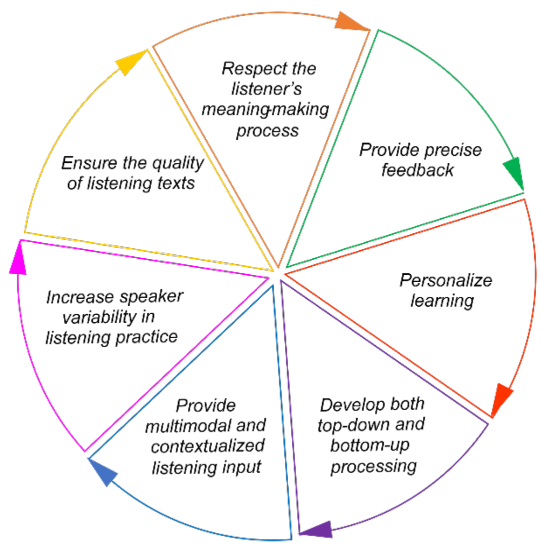
As artificial intelligence (AI) is becoming a global trend in almost every aspect of modern life, L2 researchers and practitioners are beginning to look for possible uses of AI-based software for L2 listening instructions. Multimedia-based applications with the integration of AI such as YouTube, Netflix, TuneIn radio, Spotify Music have been used for self-learning of listening comprehension (Suryana et al., 2020). Other virtual personal assistant applications such as Siri or Alexa powered by AI speech recognition and speech generation have also been used for teaching both English speaking and listening (Dizon, 2020). It is worth noting that while applications such as YouTube or Netflix are powered by AI, those applications are not developed for teaching or learning listening comprehension, and thus the AI in those apps only serves to personalize the users’ listening preferences and has nothing to do with assisting comprehension. Likewise, application like Alexa might be used for teaching and learning speaking or listening to some extent, however, they are not designed for those purposes, and thus they possess no features that support learning. Dizon (2020) pointed out in his study that the students' listening comprehension did not improve significantly after the training with the virtual personal assistant Alexa, which is understandable since the app was not designed for that purpose. Therefore, as stated earlier, intelligent CALL applications specifically developed for L2 listening comprehension are still very rare and this area of research should need more investigations.

The Learning Environment: Listening Hacked

Theoretical underpinnings

This section presents the seven principles on which the development of the proposed ICALL environment was based (see Figure 2). These principles derive from the study's theoretical framework which was constructed on radical constructivism (von Glasersfeld, 2007), precision language education (Lian & Sangarun, 2017), structure-building model of comprehension (Gernsbacher, 1990), interactive model of listening (Rost, 2015), and cognitive theory of multimedia learning (Mayer, 2009). In discussing these principles, we also explicate why certain features should be included in the learning environment.

Figure 2
Seven principles for designing CALL applications for L2 listening



(1) Respect the listener's meaning-making process

According to radical constructivism, knowledge is actively constructed by the learner and it cannot be imposed intact on them (Bodner, 1986; von Glasersfeld, 1991, 2007). Therefore, the learning environment should respect the listener's meaning-making process.

(2) Provide precise feedback

Given learning is an adaptive process and knowledge construction must have viability (von Glasersfeld, 2007), feedback provision is critical in the learning process as it is a form of awareness-raising. The effectiveness of corrective feedback (CF) depends on several factors, including, but not restricted to, the explicitness of CF (Ellis, 2017; Heift & Hegelheimer, 2017; S. Li, 2017; Nassaji & Kartchava, 2017; Quinn, 2014; Suzuki et al., 2019; Yilmaz & Granena, 2021), the immediateness of CF provision (Canals et al., 2020; Henderson, 2021; Lyster & Saito, 2010), the degree of precision and personalization of CF (Han, 2008; Pérez-Segura et al., 2020), and the reliability of CF given that it is generated by a computer (James, 2006; Li, Link, Ma, Yang, & Hegelheimer, 2014). Taking these issues into consideration, the second principle is that feedback provision should be explicit, immediate, and personalized, by targeting specific and real problems of an individual listener.

(3) Personalize learning

As discussed, personalization is an effective approach to tackling individual differences in language learning (Lian & Sangarun, 2017). The learning environment, therefore, should offer various help options or learning paths to cater for different learners effectively.

(4) Develop both top-down and bottom-up processing

In the model of listening comprehension presented earlier, listening is an interactive process between top-down and bottom-up processing, therefore, listening practice should promote the types of activity that can tap into those two primary processes as an attempt to reflect its operation in real-life situations, and with that, listening practice can better prepare learners in real L2 listening encounters.

(5) Provide multimodal and contextualized listening input

According to the cognitive theory of multimedia learning (Mayer, 2009), learners will benefit more when the learning material is presented with multimodalities (e.g. both auditory channel and visual channel) or multimedia (e.g. both pictures and written texts) than being restricted to a single medium or input modality. On that account, the learning environment needs to provide contextualized, multimodal input for the learners.

(6) Increase speaker variability in listening practice

In the model of listening comprehension, perceptual information is an inextricable source that the comprehension process must draw on. Unfortunately, the speech signal presents a great amount of variability originating from speakers' voices, allophonic variations, co-articulation, and other simplification processes in connected speech. Moreover, for biological reasons, human articulation of the same speech elements is far from constant. This variability poses a major challenge to non-native listeners' perceptual processing and thus directly affects L2 listening comprehension. A regular way of thinking to overcome this challenge is reducing the variability or simplifying the input, yet this method has not been proven empirically. On the other hand, mounting evidence in speech perception research has shown that high variability perceptual training can improve speech perception, including phoneme discrimination, word recognition, and indexical feature identification (Barriuso & Hayes-harb, 2018; Bradlow et al., 1997; Logan et al., 1991; Perrachione et al., 2011). In addressing the issue of variability in the speech signal, L2 listening instruction therefore should increase speaker variability in the input rather than reduce it, to facilitate perceptual learning (Lively et al., 1993; Logan et al., 1991).

(7) Ensure the quality of listening texts

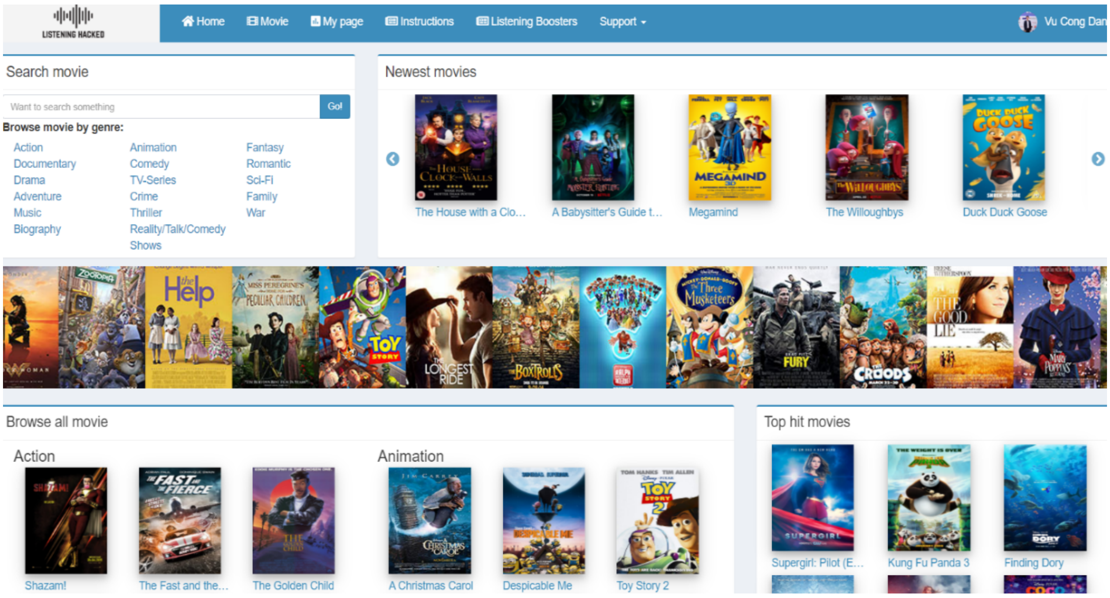
As Rost (2006) suggests, the quality of listening texts rests on the level of relevance, difficulty, and authenticity. Relevance requires listening materials to have motivational effects on the listener. In other words, the listener should have an authentic need to process and understand the listening texts. This is because real-life listening is characterized not only by cognitive processes but can be affected by the listener's willingness to listen (Worthington & Bodie, 2017). Weaver was the first listening scholar to notice the affective dimension of listening in communication when proposing in his book that "willingness to listen is probably as important as the capacity to listen" (1972, p. 8). To rephrase it, the choice that a listener makes whether to attend to the speech or to ignore it is a vital component and prerequisite of listening. Authenticity involves using genuine target-language texts whenever possible. As far as the goal of L2 listeners is to understand the language actually used by native speakers, they need to be introduced to that type of language use (Rost, 2006). Text difficulty refers to the cognitive load of a text imposing on the listening process. Factors affecting the difficulty of listening texts include "length, speed, familiarity, information density, and text organization" (Rost, 2006, p. 50). The learning environment, thus, should provide a wide range of authentic materials so that individual learners could find appropriate input for their learning.

Descriptions of the learning environment

The ICALL application presented here was designed as a web-based application and named *Listening Hacked* (see Figure 3). For the purpose of tracking the students' progress, Listening Hacked required each student to log in the website in order to use it. Additionally, a page was set up to give students detailed instructions on how to use the app in their learning.

Figure 3

A screenshot of Listening Hacked



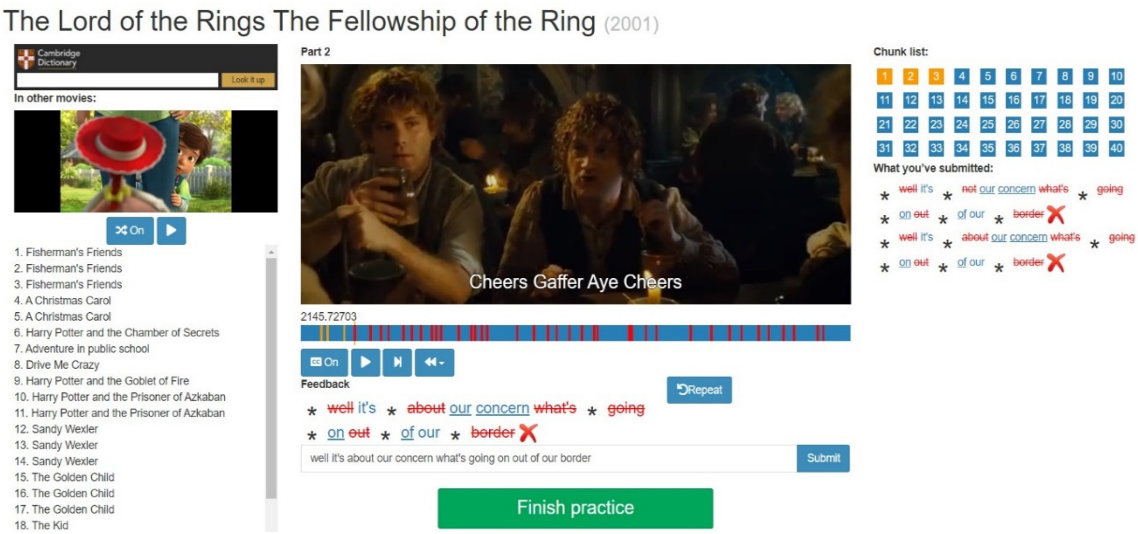
In essence, the learning on Listening Hacked involved watching English-speaking movies, with or without English captions, and occasionally transcribing some utterances from the movies that had been automatically selected by the app. The app offered a plethora of movies of various genres and classified movies according to their most relevant genres, the information of which was obtained from the Internet Movie Database (www.imdb.com/). Since this app was only used for research purposes and could be accessed by a small group of students in a specified timeframe of the research project, the use of movies in this project was legal according to Articles 25 and 32 in Vietnam's intellectual property law (Vietnam National Assembly, 2009).

A movie normally lasts from 90 to 120 minutes, thus, for convenience in practice, each movie was divided into three to four practice sessions, i.e., approximately 25 to 30 minutes per session. Students were free to choose a movie of their interest to view and practice; however, animated movies and documentaries were recommended at the outset since these movies are generally slow-paced and the language is less idiomatic as opposed to other movie genres. It remained, however, the student's own decision whether to follow this recommendation.

Paused transcription

As the students viewed the movies on Listening Hacked, they had to do paused transcription tasks (see Figure 4). For every practice session, the app automatically selected 10 chunks and removed the caption texts of those chunks. When the movie playback reached a selected chunk, the app executed a forced pause and required the students to provide transcription of the chunk immediately preceding the pause. The app did not allow the students to continue the playback unless the correct transcription was submitted, or unless they chose to skip transcribing the current chunk. The skip option became available only after five attempts by the learner.

Figure 4
Main working space



Repeated listening

Because of issues related to working memory capacity, the listener might not remember precisely what they have heard or may fail to recall some words when they encounter a pause. For this reason, the students were allowed to rehear the chunk being transcribed as many times as they wanted to with the Repeat or Rewind options. The Repeat button allowed the students to rehear the exact chunk being transcribed, while the Rewind button allowed them to rehear that chunk with more contexts, i.e. rehear the chunk plus 2 seconds before and after it. They could also choose to rewind 1, 2, 5, or 10 minutes before the chunk for more contexts. These help options respond particularly to principle (3) of the learning environment, i.e. creating personalization as a solution to individual differences in learning. Note that the unit of listening practice in this study is sometimes referred to as a chunk which is operationalized as an intonation unit (or pause unit) consisting of phrases or clauses and uttered in a single short burst of speech (Rost, 2015). In a movie, it regularly coincides with a speaker's complete utterance.

Corrective feedback

To maximize the effectiveness of repeated listening help options, the students should know exactly where to attend to; therefore, they need corrective feedback as stated in principle (2). Listening Hacked utilized a built-in program to indicate any errors in the students' transcription input for this purpose. The corrective feedback was given immediately after the students submitted their transcription texts. The mark-up software, which was based on the markup models developed by Cryle and Lian (1985) and Lian and Sangarun (2017), employed a string-matching technique that compared the student's submitted transcription with a pre-stored corresponding transcription. This technique can produce very high accuracy when capitalization and punctuation are not considered as errors. The mark-up rules are as follows.

- Word(s) which actually appear(s) in the original caption text and is(are) displayed in their correct position and will be presented in blue.
- Word(s) which actually appear(s) in the original caption text and is(are) displayed in the wrong position will be presented in blue and underlined. The feedback will indicate the correct order of words only when the students have made at least three attempts to respond.
- Word(s) which do(es) not exist in the original text will be presented in red and strikethrough.
- Word(s) which exist(s) in the original text but is(are) not transcribed will be displayed with the symbol * in their correct position.
- Mismatched punctuation and capitalization are not counted as errors and are left untouched.

Table 1 shows an example of how the mark-up rules are applied to generate feedback for the students' submitted transcriptions.

Table 1
An illustration of the mark-up rules

Submission	**Original text**	**Student's transcription**	**Feedback**
1st time	What's going on here?	What's on going?	What's * on * going *?
2nd time	What's going on here?	What's on going there?	What's * on * going * ~~there~~?
3rd time	What's going on here?	What's on going here?	What's * on$_3$ * going$_2$ here?

High variability perceptual training

When students made a transcription error, this indicated that they were experiencing a perceptual problem. In conformity with principle (6), high variability perceptual training is an effective method for overcoming this issue. The app had a feature, *Phrase Search*, for providing personalized, high variability perceptual training for any perceptual problem being experienced in the transcription tasks. Specifically, when the students clicked on any asterisk symbols in the corrective feedback, the app would execute three operations in the following order: (1) tagging part-of-speech (POS) for each word in the utterance being transcribed, (2) chunking the tagged tokens, and (3) searching the target word/chunk in the movie data.

The students were encouraged to watch all the suggested video supercuts, for the purpose of increasing exposure to various contextualized occurrences of the target word/chunk. Repeated listening to the target word/chunk in various contexts is actually a modified version of high variability training for a certain perceptual problem in listening. Note that the present study has developed a novel method for delivering high variability training. While the conventional training is aimed at phonetic contrasts by having students expose to decontextualized audio-only input (Lively et al., 1993), the type of perceptual training in this study was targeted at word and phrase level, and utilized audio-visual, contextualized materials. The modifications in the training were based on the findings of previous research on this matter, in which the inclusion of visual and contextualized information is expected to boost the effectiveness of the training (Hardison, 2003).

Other help options

The students could review their previous submissions through the list of recently submitted transcriptions. They could also choose to skip transcribing a chunk to continue viewing the movie, however, this option was only available after at least five attempts of failing to provide an accurate transcription. Any skipped word/phrase was saved in *the list of skipped items* for further practice. Additionally, the students could look up any word in the integrated dictionary.

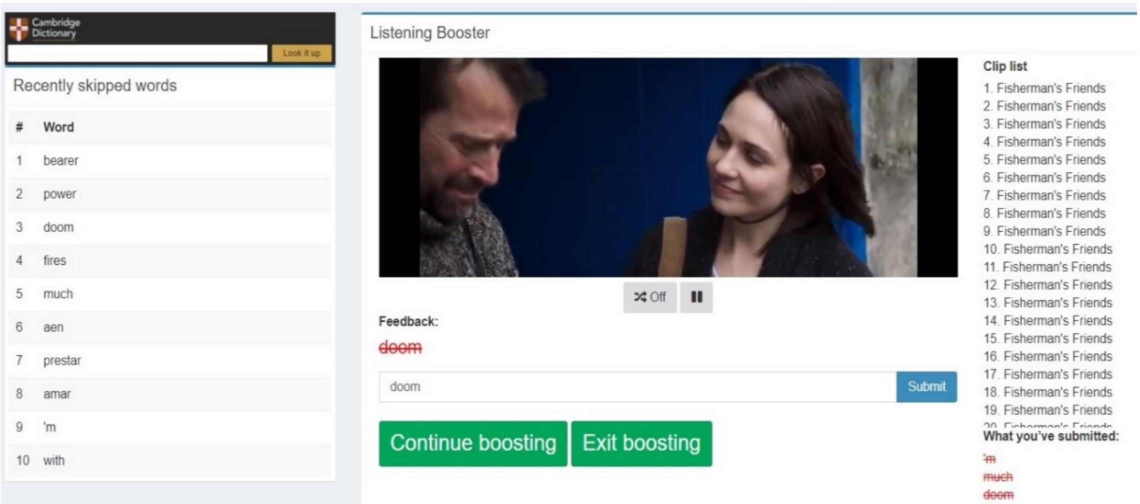
Task performance scores

When the student completed a practice session, Listening Hacked automatically calculated and displayed the task performance score for the student. *Task performance scores* (TPS) were calculated by dividing the number of successfully transcribed words by the total number of words to be transcribed and then multiplied it by 10. For example, if the student successfully transcribed 190 words out of 400 words in a practice session, the TPS was 4.75 (out of 10).

Personalized follow-up practice: Listening Boosters

Listening Hacked also had a personal page, called My Page, for showing students the summaries of what they had done in the practice sessions including their practice results, a list of skipped words, and a list of recently repaired words/phrases. By visiting their personal page, the students could have more personalized practice with *Listening Boosters* (see Figure 5). Listening Boosters generated appropriate exercises for individual students based on the list of skipped items. Whenever students clicked on the request button '*Suggest videos for me now*', the app randomly selected a word/phrase from the list of skipped items as a keyword for performing a word-in-video search, identical to the technique used in the main practice sessions. The app then presented all the video clips containing the target word/phrase and requested the students to supply just the word/phrase occurring in all the presented video clips. The app also provided corrective feedback if the students' transcription was inaccurate as well as allowing the students to submit multiple revisions. When the students could provide an accurate answer, the app removed the target word/phrase from the list of skipped items.

Figure 5
Main working space



Evaluation of the Learning Environment: Methodology

Participants

To evaluate the effectiveness of the learning environment, an experiment was conducted for 12 weeks. The study employed the simple random sampling method in which we approached all first-year, English major students in a university in Vietnam by delivering brochures, sending out emails, posting on the Facebook page of the faculty of foreign languages, and meeting the students in their classes. We randomly selected 100 students from the list of 151 students who registered to participate in the study. We then randomly assigned the 100 selected participants into the Experimental group (CG) and Control group (CG). However, the final sample of the present study was only 53 participants since many students dropped out of the study at some stage. The study took place when schools in Vietnam resumed after a long lockdown due to the Covid-19 pandemic. The students were thus overwhelmed with the pending assignments and tests, and they had to attend classes double the normal time. For this reason, many of the students decided to withdraw from their studies to focus on their regular learning.

The EG studied EFL listening on Listening Hacked and was required to complete 3 practice sessions per week, i.e around 2.25 hours per week. The CG, on the other hand, performed weekly EFL listening exercises assigned on Google Classroom. These exercises were selected from two EFL course books (Keynotes Upper-intermediate and Perspectives 3), and English podcasts accompanied with comprehension questions. They had to complete five to six exercises per week, i.e. approximately 2.25 hours per week in 12 consecutive weeks, in order to ensure an equal amount of time spent on tasks between the two groups. The experiment was designed independently of the student's learning at the university and was not part of their regular courses.

Instruments

The study used the Cambridge First Certificate (FCE) listening tests for measuring the students' pre-training and post-training levels of EFL listening (*B2 First General*, 2021). FCE listening tests are designed to assess English proficiency for general purposes and are not specifically designed for academic purposes. This made them suitable for use in the present study where the students' listening practice primarily involved viewing movies.

In addition to the pretest and posttest, the EG students' EFL listening performance was also measured through the task performance scores (TPS) which they received upon completion of a practice session. To ensure the reliability of TPS, not all TPS were collected but only those resulting from the three compulsory practice sessions which were carefully designed with the following criteria. First, the three practice sessions were compulsory for all EG students and were assigned in three designated timepoints in the research: Weeks 2, 7, and 12. Second, the sessions were taken from the first three episodes of the same movie series, Supergirl Season 1 (2015), therefore, having the same baseline story and length. Third, the 10 chunks selected for each practice session needed to satisfy the following criteria: (1) its vocabulary is profiled at B2 level, (2) the speech rate ranges from 2.6 to 3.5 words per second, or from 160 to 210 words per minute, since the average speech rate in English is 180 words per minute (Szarkowska & Gerber-Morón, 2018), (3) each chunk consists of exactly 10 words, and (4) the words selected for transcribing are not proper names.

Data analysis

Data collected from the pretest, posttest, and TPS were analyzed statistically using IBM SPSS Statistics (Version 23) to determine the effects of Listening Hacked on the students' EFL listening ability. The analytical procedures involved (1) comparing the differences between the pretest and posttest scores within each group, (2) comparing the differences in TPS in the three observed timepoints, and (3) comparing the posttest performances between the two groups.

Procedures for a practice session

The students were free to choose a movie of their interest to start a practice session. As they watched the movie, they had to transcribe a total of 10 randomly selected chunks from the movie at random intervals to continue the movie playback. There was no restriction on the time and number of attempts that they spent on transcribing each chunk. They could also use the help options as much as they desired. The help options, as described earlier, include Repeat, Rewind, Corrective feedback, Phrase search, and a built-in dictionary.

The students could skip a chunk during a practice session and continue the movie playback only after they had submitted at least five attempts. A practice session was recorded as complete when all the selected chunks were solved (or skipped). Students could watch a movie again in another practice session; however, new chunks could be selected by the app.

Results

Pre and post-training differences in EFL listening performance

The first objective of the analyses was to examine whether there were any differences between the students' pre and post-training ability in EFL listening comprehension. These analyses using paired samples t-test were applied to both EG and CG.

The results of the paired samples t-test on the pretest and posttest (

Table) show that the EG's listening test scores improved by 3.30 on average (N=30; t=4.134; p<.01), with a relatively large effect size (d=.76) while the CG's pretest and posttest scores did not differ significantly (N=23; t=1.416; p>.05). Since there might be concerns over the small sample size of the CG, the bootstrap method was invoked, based on 1000 resampling times. The bootstrap result (Table) shows that the 95% CI includes value zero (Lower bound: .-348; Upper bound: 2.43), indicating a statistical lack of significance. This is in line with the paired samples t-test results.

Table 2
Paired samples t-test on the Pretest and Posttest

		Experimental Group	**Control Group**
		Mean difference (Posttest – Pretest)	Mean difference (Posttest – Pretest)
Paired Differences	Mean	3.3	1.04
	Std. Deviation	4.37	3.54
	Std. Error Mean	0.798	0.737
T		4.134	1.416
Df		29	22
P-value (2-tailed)		.000	0.171

Table 3
Bootstrap results of the paired samples t-test on the CG Pretest and Posttest

	Mean	**Bootstrap**				
		Bias	Std. Error	P-value (2-tailed)	95% Confidence Interval	
					Lower	Upper
Posttest – Pretest (CG)	1.04348	0.00783	0.72159	0.173	-0.34783	2.43478

To put it concisely, the students in the EG improved their EFL listening comprehension performance significantly after the training with Listening Hacked, while the students in the CG did not. Interestingly, 30.43% of the CG performed even worse in the posttest than in the pretest while this proportion was lower in the EG (23.33%).

Differences in EFL listening performance between three observed timepoints

In addition to the pretest and posttest, another data source of listening performance was the task performance scores (TPS) which the EG students received upon completing the compulsory practice sessions in the three timepoints: Weeks 2, 7, and 12. Week 2 was selected because that was the time the students started using Listening Hacked and the other two weeks were chosen to ensure the equal intervals. The Mauchly's test of sphericity indicates that the dataset does not violate the assumption of sphericity ($\chi2(2)$=.876; p>.05), thus the ANOVA results should be interpreted with sphericity assumed. The repeated measures ANOVA with sphericity assumed (Table) reveals that the differences in the TPS between the three timepoints were statistically significant (F(2, 58)=22.373; p<.001).

Table 4
Repeated measures of ANOVA

Source		df	F	P-value
Week	Sphericity-assumed	2	22.373	.000
Error (Week)	Sphericity-assumed	58		

The post-hoc tests using Bonferroni correction (Table) were also conducted to investigate further where the differences occurred. The results show that the differences were statistically significant between Week 2 and Week 7 (mean difference=-1.507; p<.001), and between Week 2 and Week 12 (mean difference=-1.885; p<.001). There was a difference of -.378 between Week 7 and Week 12, however, it did not reach a statistical significance level at .05. In short, the results suggest that the students' TPS increased significantly over the observed timepoints.

Table 5
Post-hoc tests for pair-wise comparisons

Week (I)	Week (J)	Mean Difference (I-J)	Std. Error	P-value
2	7	-1.507	.254	.000
	12	-1.885	.334	.000
7	12	-.378	.289	.000

Post-training differences in EFL listening performances between EG and CG

To examine the comparative effects of the training using Listening Hacked on the students' EFL listening, an independent samples t-test was run on the posttest. The results (Table) show that the EG outperformed the CG significantly on the posttest, with a mean difference of 4.24 (t=2.861; p<.05). Since equal variance was assumed in the data according to the Levene's test result, the assumption of normal distribution was not a concern in this analytical procedure. However, to corroborate the results of the t-test, the bootstrap method was still performed for calculating the 95% CI. The results of bootstrap for independent samples t-test (Table) show that the 95% CI does not include value zero (Lower bound 1.60; Upper bound 6.82) which indicates a similar result as compared to the t-test results.

Table 6
Independent t-test on the Posttest

Levene's Test for Equality of Variances		**t-test for Equality of Means**				
F	P-value	t	df	P-value (2-tailed)	Mean Difference	Std. Error Difference
3.087	0.085	2.861	51	0.006	4.23913	1.48183

Table 7
Bootstrap results of the independent t-test on the Posttest

Mean Difference	**Bootstrap**				
	Bias	Std. Error	P-value (2-tailed)	95% Confidence Interval	
				Lower	Upper
4.23913	0.04926	1.36715	0.005	1.59506	6.81685

Discussion

The effectiveness of the learning system

The present research set out to investigate the effectiveness of the proposed learning platform, Listening Hacked, on the students' EFL listening comprehension. The results show that after 12 weeks of learning on Listening Hacked, the students in the EG improved their EFL listening performance significantly, with an increase of 11% and a relatively large effect size (d=.76), while the students in the CG who were exposed to a traditional method did not show evidence of improvement. In fact, 30.43% of the students in the CG deteriorated in their posttest performance as compared to 23.33% of the EG. The EG's improvement and outperformance in EFL listening comprehension suggest that Listening Hacked was effective in assisting the students in their learning. The effectiveness of Listening Hacked could possibly be attributed to the student's development in perceptual processing and learner autonomy in learning EFL listening.

Improvement in perceptual processing

Previous studies show that improvement in perceptual processing could lead to better comprehension in L2 listening (Field, 2003; Jia & Hew, 2019; Leonard, 2019; Matthews & O'Toole, 2015). The present study seems to suggest similar conclusions. In this study, the students' improved performance in paused transcription tasks as evidenced by the significant increases in the task performance scores (TPS) implies that their word recognition ability had improved and thus led to better comprehension in EFL listening. Field (2003) asserts that many problems in listening, including those related to higher level processes, are in fact rooted in problems of perceptual decoding. Without adequate recognizable perceptual information, top-down information would become useless, and the speech would become unintelligible. In a study examining learners' problems in L2 listening, Goh (2000) found that several listening problems were related to the learners' inability to recognize words. For example, they could not recall the meanings of words in speech although they said that those words were familiar to them. This happened because the learners could not establish the link between the spoken and the written forms of the words. This link is established through the same linguistic representations to which both spoken and written forms need to contact in speech recognition (Johnsrude & Buchsbaum, 2016). While the connection between the written forms and their corresponding representations in the brain is normally robust, that between the spoken forms and the linguistic representations may not always be easily established since the spoken form of a word can have multiple variants. This explains why one may find some words familiar in some cases and the same words unrecognizable in other cases. In this perspective, the paused transcription tasks in this study directly addressed this issue. The tasks provided the students with opportunities to practice establishing the link between written and spoken forms of words and overcoming the problems of variability in the speech signal. As a result, the student's perceptual decoding ability improved and their listening comprehension followed.

Attention problems in L2 listening could also stem from problems in word recognition. For example, some learners in Goh's (2000) study said that they missed the next part of the text when spending time thinking about the meanings of certain words. Due to the limited capacity of working memory (Baddeley, 1992, 2012), more attention allotted to perceptual processing means that less attentional resources are spared for other processes, and that would also affect the perceptual processing of subsequent speech signals. According to the view that working memory is a dual function system consisting of processing and storage (Daneman, 1991; Finardi, 2006), when the processing function does not take much space, more space is available for storage, and the limited capacity of working memory is well used to process and store more incoming information. In other words, it could be said that the study helped to improve the students' working memory by reducing the negative impacts of the trade-off between perceptual processing efficacy and allocation of attentional resources. Therefore, improvement in perceptual processing or word recognition ability could help solve attention problems in listening and positively affect general EFL listening ability.

Development in learner autonomy

The students' improvement in EFL listening comprehension could also be explained by the development in learner autonomy. While autonomy was not measured explicitly in this study, it was an obvious part of learning on Listening Hacked which is an autonomous learning environment. Unlike the CG studying with a traditional approach, the students in the EG had to decide almost everything in their learning, including choosing the appropriate listening materials, developing a set of strategies for using help options, and working on their own listening problems. Those decisions and learning behaviors clearly contributed to their growth in autonomy as the construct of learner autonomy is defined as the ability to take responsibility for one's learning (Holec, 1981). Research shows that learner autonomy is actually correlated with EFL proficiency (Dafei, 2007) and EFL listening comprehension ability (Safari & Tabatabaei, 2016). Therefore, the students' development in learner autonomy could also lead to better learning, hence better EFL listening ability.

The slowing down of listening development

The repeated measures ANOVA on the task performance scores (TPS) over the two time periods in the experiment reveals some interesting facts: the TPS in Week 7 increased 15% compared to that in Week 2, while the TPS in Week 12 was only 3.8% higher than that in Week 7. This implies that the students' improvement in EFL listening ability apparently slowed down with time. One possibility is that the learning platform has its own limits, and hence it might not be capable of helping the students grow further once they reach a certain threshold. This could happen when they have used up the available help options in the platform and none of those could bring the necessary support for them to overcome certain listening problems. In this case, it is important that the student's performance data be obtained and analyzed so that new functionalities can be devised accordingly.

Another possible explanation is that higher levels of listening proficiency may require longer practice time than lower levels. That is to say, an increase of 2.0 TPS from 2.0 to 4.0 could be less effortful and take much less time than that of 2.0 TPS from 8.0 to 10 (the maximum TPS). According to a Cambridge University Press report, to achieve A1 level in the CEFR, an adult learner at beginner's level would need to spend 90 to 100 hours on guided learning, whereas to reach C1 level from B2 level, they should spend 200 to 300 hours of learning, and an additional of 300 to 400 hours for achieving C2 level (Knight, 2018). This means that the speed of improvement did not slow down, however, higher levels of listening proficiency encompass more requirements, both in linguistic knowledge and processing skills, and thus they simply demanded more time. If that were the case, the implication here is that the learning platform should be frequently updated to provide adequate and various learning resources for extensive listening practice in order to ensure the students' continuing development.

Implications of the Study

With the encouraging results of the implementation of the proposed technology-based learning platform, we discuss some implications of the study below.

First, the results of this study suggest that personalization in learning can be achieved by incorporating appropriate technology. The construction of the multimedia database in this study illustrates how technology can allow us to create a large collection of listening materials that suits every learner in the sense that learners are offered more learning choices and can decide what fits their learning preferences and experience. These conceivable virtues of technology can solve the matter of input inappropriacy caused by the conventional classrooms in which "the whole class" is obliged to study the same materials at the same time, in the same way and in the same space. The database could be used as main listening materials or as a reference library like the Listening Boosters feature. Alternatively, cooperative efforts could be launched to combine the efforts of teachers and researchers to produce high quality shareable databases for use by all learners. Such a development would maximize the developers' potential and would provide greater benefit to the learner community.

Second, this study highlights the capabilities of technology for input enhancement (Cardenas-Claros & Gruba, 2013). These capacities are associated with (a) media playback controls (e.g. play/pause/rewind buttons) and (b) multimedia materials (e.g. videos, captions) (Hubbard, 2017). Regarding the ephemeral nature of listening, using media players with playback controls for listening practice can help learners pace their own learning so that perceptual processing overload is efficiently addressed on an individual basis. Media controller bars and sliders, for example, or the Repeat/Rewind buttons can help increase the accessibility of authentic materials by enabling students to rewind the speech to a certain part and listen again to it as many times as they wish. Note that in order to optimize the usefulness of sliders, we should place certain restrictions in their usage, that is, they might allow students to rewind freely, but not permit them to fast-forward or skip over the listening texts. Alternatively, deficiency in comprehension can be tackled when learners are given options for using multimedia such as videos (Kim, 2015) or captioned videos (Gass, Winke, Isbell, & Ahn, 2019; Winke, Gass, & Sydorenko, 2010) in their listening practice. The study, therefore, suggests that videos be used both as the primary listening materials as in this study and as feedback or references in L2 listening teaching and learning.

Third, another implication of this study is the possibility that technology-based autonomous learning environments like Listening Hacked could make L2 listening learning effective without the presence of the teacher while, at the same time, being capable of solving the paradox between the large-sized language classroom and the teacher's inability to cater for individual learners' needs. In the future, when more ICALL apps are developed, online teacherless classrooms might be a practical solution to L2 listening learning. As this study took place during the COVID-19 pandemic with social distancing policies, it further demonstrates that autonomous learning platforms as such may be an indispensable part of L2 listening instruction in the future and of L2 pedagogy in general.

Finally, many students find listening to a foreign language difficult because they cannot recognize the spoken words in connected speech. This study has demonstrated that high variability perceptual learning is not only effective in learning phonetic contrasts (Logan et al., 1991) but useful in training word recognition and phrasal recognition as well. This is a promising alternative way of addressing perceptual problems in L2 listening. Furthermore, the use of audio-visual materials has shown value as compared to the traditional audio-only materials in high variability phonetic training. The study also implies that perceptual training should be personalized and not be based on a predefined list of overgeneralized perceptual problems among L2 learners. That means the learner should only work on their own perceptual problems as they feel necessary, not the common problems of the group they belong to.

Limitations and future work

Although the study has produced valuable insights and implications for researching and teaching L2 listening, it cannot avoid certain limitations.

Firstly, the study had a relatively small sample size and took its samples from only one university. In addition, the participants in this study were English-major undergraduates. While this sampling technique was useful for producing a homogeneous group and advantageous to statistical analyses, generalizations to other contexts and non-English major students might be cautious. This study could be replicated using a larger sample size and include a variety of participants whose levels of EFL listening proficiency might be different.

Secondly, as the present study investigated the effectiveness of the app in 12 weeks, further work could assess the long-term effects on the students' EFL listening comprehension. This will help to explore the limits of the methods used in the app in developing students' listening ability. On the other hand, as the students' listening development in this study was slowed down in the last 5 weeks of the experiment, further research could also investigate the optimal time of using the app as to maximize its effectiveness in the shortest time.

Thirdly, since the present study only examined the overall effects of Listening Hacked on the students' EFL listening comprehension, the effects of the separate help options and instructional methods embedded in the design of the app have not been measured. Therefore, the question of to what extent a certain help option or a pedagogic decision underlying the design of Listening Hacked influenced the students' EFL listening comprehension remains unknown. It is also difficult to conclude whether the use of automated corrective feedback, for example, was contributing to the students' listening improvement more or less than that of Phrase Search. Therefore, further work needs to be done to determine the effects of separate methods employed in this study on L2 listening comprehension by isolating the effects of each method or help option. A possible way is to include more groups of participants in the research design, and each group will be allowed to use only one method or help option in their learning. Such investigations will allow the comparisons of certain help options in assisting L2 listening learning.

Conclusion

This study investigated the effectiveness of an original learning platform constructed on the proposed framework for designing CALL applications for L2 listening. The statistical analyses show that the students in the Experimental Group significantly improved their EFL listening performance on the posttest after 12 weeks of learning on the platform. They also outperformed their counterparts who did traditional listening exercises on Google Classroom. This improvement in EFL listening was indicative of the student's development of spoken word recognition ability and learner autonomy. The study thus offers insights into how NLP and multimedia databases could be integrated into CALL applications for achieving the goal of precision language education, facilitating the learning process of L2 listening, and at the same time promoting learner autonomy. In this study, we have discussed and demonstrated that with a principled framework, the development and implementation of technology-based solutions like Listening Hacked are feasible and could produce good results. However, much work still needs to be done in research and development to produce and implement more successful

ICALL applications for L2 listening. Researchers and practitioners can consider this study as a guideline and a source of inspiration for taking further steps in creating comprehensive and practical solutions for L2 listening learning and teaching.

References

Al-Athwary, A. A. H., & Lasloum, N. M. (2021). Second Language Listening Comprehension Gain From Aural vs. Audio-Visual Inputs: The Case of EFL Arab Learners. *Journal of Language Teaching and Research*, *12*(6), 911–921. https://doi.org/10.17507/jltr.1206.07

Azizinezhad, M., & Hashemi, M. (2013). A Look at the Status of Computer Assisted Language Learning and its Applications. *Procedia - Social and Behavioral Sciences*, *93*, 121–124. https://doi.org/10.1016/j.sbspro.2013.09.163

B2 First General. (2021). https://www.cambridgeenglish.org/exams-and-tests/first/exam-format/

Baddeley, A. (1992). Working Memory: The Interface between Memory and Cognition. *Journal of Cognitive Neuroscience*, *4*(3), 281–288. https://doi.org/10.1162/jocn.1992.4.3.281

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Barriuso, T. A., & Hayes-harb, R. (2018). High Variability Phonetic Training as a Bridge From Research to Practice. *The CATESOL Journal*, *30*(1), 177–194. http://www.catesoljournal.org/wp-content/uploads/2018/03/CJ30.1_barriuso.pdf

Benson, P. (2013). Teaching and Researching: Autonomy in Language Learning. In *Teaching and Researching: Autonomy in Language Learning* (2nd ed.). Routledge. https://doi.org/10.4324/9781315833767

Bhutoria, A. (2022). Personalized education and Artificial Intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop model. *Computers and Education: Artificial Intelligence*, *3*. https://doi.org/10.1016/j.caeai.2022.100068

Bodner, G. M. (1986). Constructivism: A theory of knowledge. *Journal of Chemical Education*, *63*, 873–878. http://dx.doi.org/10.1021/ed063p873

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English / r / and / l /: IV. Some effects of perceptual learning on speech production . *The Journal of the Acoustical Society of America*, *101*(4), 2299–2310. https://doi.org/10.1121/1.418276

Brett, P. (1995). Multimedia for listening comprehension: The design of a multimedia-based resource for developing listening skills. *System*, *23*(1), 77–85. https://doi.org/10.1016/0346-251X(94)00054-A

Brodbeck, C., Bhattasali, S., Cruz Heredia, A. A. L., Resnik, P., Simon, J. Z., & Lau, E. (2022). Parallel processing in speech perception with local and global representations of linguistic context. *ELife*, *11*. https://doi.org/10.7554/eLife.72056

Canals, L., Granena, G., Yilmaz, Y., & Malicka, A. (2020). Second Language Learners' and Teachers' Perceptions of Delayed Immediate Corrective Feedback in an Asynchronous Online Setting An Exploratory Study. *TESL Canada Journal*, *37*(2), 181–209. https://doi.org/10.18806/tesl.v37i2.1336

Cardenas-Claros, M. S., & Gruba, P. A. (2013). Decoding the "CoDe": a framework for conceptualizing and designing help options in computer-based second language listening. *ReCALL*, *25*(2), 250–271. https://doi.org/10.1017/S0958344013000049

Cárdenas-Claros, M. S., & Gruba, P. A. (2009). Help Options in CALL: A Systematic Review. *CALICO Journal*, *27*(1), 69–90. https://doi.org/10.11139/cj.27.1.69-90

Cárdenas-Claros, M. S., & Gruba, P. A. (2014). Listeners' interactions with help options in CALL. *Computer Assisted Language Learning*, *27*(3), 228–245. https://doi.org/10.1080/09588221.2012.724425

Çetinkaya, I. G., & Tilfarlioglu, F. Y. (2020). Three factors affecting language learning: Grammar learning strategies, self-efficacy, and learner autonomy. *Universal Journal of Educational Research*, *8*(7), 2929–2936. https://doi.org/10.13189/ujer.2020.080721

Chi, C., Chen, H.-J. H., Tseng, W.-T., & Liu, Y.-T. (2022). Efficacy of different presentation modes for L2 video comprehension: Full versus partial display of verbal and nonverbal input. *ReCALL*, 1–17. https://doi.org/10.1017/s0958344022000088

Cook, C. R., Kilgus, S. P., & Burns, M. K. (2018). Advancing the science and practice of precision education to enhance student outcomes. *Journal of School Psychology*, *66*, 4–10. https://doi.org/10.1016/j.jsp.2017.11.004

Cross, J. (2018). Video in Listening. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–6). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118784235.eelt0606

Cryle, P. M., & Lian, A.-P. (1985). Sorry I'll Play That Again. In J. A. Bowden & S. Lichtenstein (Eds.), *Student Control of Learning: Computers in Tertiary Education* (pp. 204–213). University of Melbourne. https://doi.org/https://doi.org/10.13140/2.1.2891.9044

Dafei, D. (2007). An exploration of the relationship between learner autonomy and English proficiency. *Asian EFL Journal*, *24*(4), 1–23. http://www.asian-efl-journal.com/pta_Nov_07_dd.pdf

Daflizar, Sulistiyo, U., & Kamil, D. (2022). Language Learning Strategies and Learner Autonomy: The Case of Indonesian Tertiary EFL Students. *LEARN Journal: Language Education and Acquisition Research Network*, *15*(1), 257–281. https://so04.tci-thaijo.org/index.php/LEARN/article/view/256724

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, *20*(6), 445–464. https://doi.org/10.1007/BF01067637

Diao, Y., Chandler, P., & Sweller, J. (2007). The effect of written text on comprehension of spoken English as a foreign language. *American Journal of Psychology*, *120*(2), 237–261. https://doi.org/10.2307/20445397

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning and Technology*, *24*(1), 16–26. https://doi.org/10.125/44705

Ellis, R. (2017). Oral corrective feedback in L2 classrooms: What We Know so Far. In N. Hossein & K. Eva (Eds.), *Corrective Feedback in Second Language Teaching and Learning: Research, Theory, Applications, Implications* (pp. 3–18). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315621432-2/oral-corrective-feedback-l2-classrooms-rod-ellis

Fan, J. (2020). Design of Hardware and Software Systems for English Listening Test

Platform. *Proceedings of 2020 IEEE International Conference on Artificial Intelligence and Information Systems, ICAIIS 2020*, 670–674. https://doi.org/10.1109/ICAIIS49377.2020.9194858

Field, J. (1999). "Bottom-up" and "top-down." *ELT Journal*, *53*(4), 338–339. https://doi.org/10.1093/eltj/53.4.338

Field, J. (2003). Promoting perception: Lexical segmentation in L2 listening. *ELT Journal*, *57*(4), 325–334. https://doi.org/10.1093/elt/57.4.325

Field, J. (2008). Guest editor's introduction Emergent and divergent: A view of second language listening research. *System*, *36*, 2–9. https://doi.org/10.1016/j.system.2008.01.001

Finardi, K. (2006). Working memory capacity and speech production in L2: evidences from a picture description task. *Rev. Est. Ling.*, *14*(1), 231–260. https://doi.org/10.17851/2237-2083.14.1.231-260

Flowerdew, J., & Miller, L. (2001). *Second Language Listening*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667244

Gass, S., Winke, P., Isbell, D. R., & Ahn, J. (2019). How captions help people learn languages: A working-memory, eye-tracking study. *Language Learning and Technology*, *23*(2), 84–104. https://doi.org/https://doi.org/10125/44684

Gernsbacher, M. A. (1990). *Language Comprehension As Structure Building*. Psychology Press. https://doi.org/10.4324/9780203772157

Gernsbacher, M. A. (1991). Cognitive Processes and Mechanisms in Language Comprehension: The Structure Building Framework. *Psychology of Learning and Motivation - Advances in Research and Theory*, *27*(C), 217–263. https://doi.org/10.1016/S0079-7421(08)60125-5

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes*, *23*(3), 265–304. https://doi.org/10.1080/01638539709544994

Gernsbacher, M. A., & Faust, M. E. (1991). The Mechanism of Suppression: A Component of General Comprehension Skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(2), 245–262. https://doi.org/10.1037/0278-7393.17.2.245

Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, *28*(1), 55–75. https://doi.org/10.1016/S0346-251X(99)00060-3

Goh, C. C. M. (2016). Cognition, Metacognition, and L2 Listening. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 214–228). Routledge. https://doi.org/10.4324/9781315716893-16

Goh, C. C. M., & Vandergrift, L. (2021). *Teaching and Learning Second Language Listening* (2nd ed.). Routledge. https://doi.org/10.4324/9780429287749

Grgurović, M., & Hegelheimer, V. (2007). Help options and multimedia listening: Students' use of subtitles and the transcript. *Language Learning and Technology*, *11*(1), 45–66. https://doi.org/10125/44088

Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. *System*, *36*(1), 85–93. https://doi.org/10.1016/j.system.2007.11.005

Han, Z. (2008). Fine-tuning Corrective Feedback. *Foreign Language Annals*, *34*(6), 582–599. https://doi.org/10.1111/j.1944-9720.2001.tb02105.x

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, *24*(4), 495–522. https://doi.org/10.1017/S0142716403000250

Hashemian, M., & Soureshjani, K. H. (2011). The interrelationship of autonomy, motivation, and academic performance of Persian L2 learners in distance education contexts. *Theory and Practice in Language Studies*, *1*(4), 319–326. https://doi.org/10.4304/tpls.1.4.319-326

Heift, T. (2012). Intelligent Computer-Assisted Language Learning. In *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd. https://doi.org/10.1002/9781405198431.wbeal0548

Heift, T., & Hegelheimer, V. (2017). Computer-assisted corrective feedback and language learning. In H. Nassaji & E. Kartchava (Eds.), *Corrective Feedback in Second Language Teaching and Learning* (pp. 51–65). Routledge. https://doi.org/10.4324/9781315621432-5

Henderson, C. (2021). The effect of feedback timing on L2 Spanish vocabulary acquisition in synchronous computer-mediated communication. *Language Teaching Research*, *25*(2), 185–208. https://doi.org/10.1177/1362168819832907

Holec, H. (1981). *Autonomy in Foreign Language Learning*. Pergamon Press. https://www.worldcat.org/title/autonomy-and-foreign-language-learning/oclc/264288592?referer=di&ht=edition

Hubbard, P. (2017). Technologies for Teaching and Learning L2 Listening. In C. A. Chapelle & S. Sauro (Eds.), *The Handbook of Technology and Second Language Teaching and Learning* (pp. 93–106). Wiley. https://doi.org/10.1002/9781118914069.ch7

IBM Corp. (2015). *IBM SPSS Statistics for Windows* (No. 23). IBM Corp.

James, C. (2006). Validating a computerized scoring system for assessing writing and placing students in composition courses. *Assessing Writing*, *11*(3), 167–178. https://www.learntechlib.org/p/103007/

Jia, C., & Hew, K. F. T. (2019). Supporting lower-level processes in EFL listening: the effect on learners' listening proficiency of a dictation program supported by a mobile instant messaging app. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2019.1671462

Johnsrude, I. S., & Buchsbaum, B. R. (2016). Speech Perception and Spoken Word Recognition. In G. Gaskell & J. Mirković (Eds.), *Speech perception and spoken word recognition*. Psychology Press. https://doi.org/10.4324/9781315772110

Kim, H.-S. (2015). Using Authentic Videos to Improve EFL Students' Listening Comprehension. *International Journal of Contents*, *11*(4), 15–24. https://doi.org/10.5392/IJoC.2015.11.4.015

Knight, B. (2018). How long does it take to learn a foreign language. In *Experience Better Learning* (Issue january 2018). Cambridge University Press. https://www.cambridge.org/elt/blog/wp-content/uploads/2018/10/How-long-does-it-take-to-learn-a-foreign-language.pdf

Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. Longman.

Lennox, R., Hepburn, K., Leaman, E., & van Houten, N. (2020). 'I'm probably just gonna skim': an assessment of undergraduate students' primary scientific literature reading approaches. *International Journal of Science Education*, *42*(9), 1409–1429. https://doi.org/10.1080/09500693.2020.1765044

Leonard, K. R. (2019). Examining the relationship between decoding and comprehension in L2 listening. *System*, *87*. https://doi.org/10.1016/j.system.2019.102150

Lesnov, R. O. (2022). Furthering the argument for visually inclusive L2 academic listening tests: The role of content-rich videos. *Studies in Educational Evaluation*, *72*. https://doi.org/10.1016/j.stueduc.2021.101087

Li, S. (2017). Student and teacher beliefs and attitudes about oral corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *Corrective Feedback in Second Language Teaching and Learning* (pp. 143–157). Routledge. https://doi.org/10.4324/9781315621432-11

Li, Z., Link, S., Ma, H., Yang, H., & Hegelheimer, V. (2014). The role of automated writing evaluation holistic scores in the ESL classroom. *System*, *44*(1), 66–78. https://doi.org/10.1016/j.system.2014.02.007

Lian, A.-P., & Sangarun, P. (2017). Precision Language Education: A Glimpse Into a Possible Future. *GEMA Online® Journal of Language Studies*, *17*(4), 1–15. https://doi.org/https://doi.org/10.17576/gema-2017-1704-01

Liang, Q. (2020). An empirical study on the effectiveness of multimedia annotation to the news listening comprehension. *Higher Education of Social Science*, *19*(1), 70–74. http://flr-journal.org/index.php/hess/article/view/11887

Little, D., Dam, L., & Legenhausen, L. (2017). Language learner autonomy: Theory, practice and research. In *Language Learner Autonomy: Theory, Practice and Research*. Multilingual Matters. https://doi.org/10.21832/LITTLE8590

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, *94*(3), 1242–1255. https://doi.org/10.1121/1.408177

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese Listeners To Identify English /R/ And /1/: A First Report. *Journal of the Acoustical Society of America*, *89*(2), 874–886. https://doi.org/10.1121/1.1894649

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*, 251–286. https://doi.org/10.1017/S0272263100009165

Luan, H., & Tsai, C. C. (2021). A Review of Using Machine Learning Approaches for Precision Education. *Educational Technology and Society*, *24*(1), 250–266. https://www.jstor.org/stable/26977871

Lyster, R., & Saito, K. (2010). Oral Feedback In Classroom SLA. *Studies in Second Language Acquisition1*, *32*, 265–302. https://doi.org/10.1017/S0272263109990520

Makhluf, H. A. (2020). Precision education: engineering learning, relevancy, mindset, and motivation in online environments. In C. M. Sistek-Chandler (Ed.), *Exploring Online Learning Through Synchronous and Asynchronous Instructional Methods* (pp. 202–224). IGI Global. https://doi.org/10.4018/978-1-7998-1622-5.ch009

Mathew, N. G., & Alidmat, A. O. H. (2013). A Study on the Usefulness of Audio-Visual Aids in EFL Classroom: Implications for Effective Instruction. *International Journal of Higher Education*, *2*(2). https://doi.org/10.5430/ijhe.v2n2p86

Matthews, J., & O'Toole, J. M. (2015). Investigating an innovative computer application to improve L2 word recognition from speech. *Computer Assisted Language Learning*, *28*(4), 364–382. https://doi.org/10.1080/09588221.2013.864315

Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

McNamara, D. S., & Magliano, J. (2009). Chapter 9 Toward a Comprehensive Model of Comprehension. In B. H. Ross (Ed.), *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 51, pp. 297–384). https://doi.org/10.1016/S0079-7421(09)51009-2

Meskill, C. (1996). Listening Skills Development Through Multimedia. *Journal of Educational Multimedia and Hypermedia*, *5*(2), 179–201. http://dcmp.org/caai/nadh166.pdf

Mompean, A. R., & Guichon, N. (2009). Assessing the use of aids for a computer-mediated task: Taking notes while listening. *The JALT CALL Journal*, *5*(2), 29–44. https://doi.org/10.29140/jaltcall.v5n2.79

Nassaji, H., & Kartchava, E. (Eds.). (2017). *Corrective Feedback in Second Language Teaching and Learning*. Routledge. https://www.routledge.com/Corrective-Feedback-in-Second-Language-Teaching-and-Learning-Research/Nassaji-Kartchava/p/book/9781138657298

Ngurah, I. P., & Myartawan, W. (2013). the Correlation Between Learner Autonomy and English Proficiency of Indonesian Efl College Learners. *TEFLIN Journal*, *24*(1), 63–81.

Nguyen, C.-D., & Newton, J. (2018). Schemata in Listening Comprehension. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–7). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118784235.eelt0592

Palfreyman, D. M., & Benson, P. (2018). Autonomy and Its Role in English Language Learning: Practice and Research. In X. Gao (Ed.), *Second Handbook of English Language Teaching* (pp. 1–21). Springer International Publishing. https://doi.org/10.1007/978-3-319-58542-0_38-1

Pangaribuan, T., Sinaga, A., & Sipayung, K. T. (2017). The Effectiveness of Multimedia Application on Students Listening Comprehension. *English Language Teaching*, *10*(12), 212. https://doi.org/10.5539/elt.v10n12p212

Pérez-Segura, J. J., Sánchez Ruiz, R., González-Calero, J. A., & Cózar-Gutiérrez, R. (2020). The effect of personalized feedback on listening and reading skills in the learning of EFL. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2019.1705354

Perrachione, T. K., Lee, J., Ha, L. Y. Y., & Wong, P. C. M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America*, *130*(1), 461–472. https://doi.org/10.1121/1.3593366

Phillips, M. D., Lowe, M. J., Lurito, J. T., Dzemidzic, M., & Mathews, V. P. (2001). Temporal lobe activation demonstrates sex-based differences during passive listening. *Radiology*, *220*(1), 202–207. https://doi.org/10.1148/radiology.220.1.r01jl34202

Pujolà, J. T. (2002). CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL*, *14*(2), 235–262. https://doi.org/10.1017/S0958344002000423

Quinn, P. (2014). *Delayed versus immediate corrective feedback on orally produced passive errors in English* [University of Toronto]. http://hdl.handle.net/1807/65728

Reinders, H. (2018). Technology and Autonomy. *The TESOL Encyclopedia of English Language Teaching*, 1–5. https://doi.org/10.1002/9781118784235.eelt0433

Rost, M. (2001). Listening. In R. Carter & D. Nunan (Eds.), *The Cambridge Guide to*

Teaching English to Speakers of Other Languages (pp. 7–13). Cambridge University Press. https://doi.org/10.1017/CBO9780511667206.002

Rost, M. (2006). Areas of research that influence L2 listening instruction. *Current Trends in the Development and Teaching of the Four Language Skills*, 47–74. https://doi.org/Book_Doi 10.1515/9783110197778

Rost, M. (2015). *Teaching and Researching Listening*. Routledge. https://doi.org/10.4324/9781315732862

Safari, S., & Tabatabaei, O. (2016). Relationship between Autonomy and Listening Comprehension Ability among Iranian EFL Learners. *International Journal of Foreign Language Teaching and Research*, *4*(13), 33–46. https://doi.org/10.5539/elt.v8n12p158

Sakai, H. (2018). Working Memory in Listening. *The TESOL Encyclopedia of English Language Teaching*, 1–6. https://doi.org/10.1002/9781118784235.eelt0579

Sakai, S., & Takagi, A. (2009). Relationship between learner autonomy and English language proficiency of Japanese learners. *The Journal of ASIA TEFL*, *6*(3), 297–325.

Sarani, A., Behtash, E. Z., & Arani, S. M. N. (2014). The Effect of Video-Based Tasks in Listening Comprehension of Iranian Pre-intermediate EFL Learners. *Gist Education and Learning Research Journal*, *8*, 29–47. https://doi.org/10.26817/16925777.112

Scovel, T. (1998). *Psycholinguistics*. Oxford University Press.

Shalmani, H. B. (2020). Does modality matter? On the modality of visuals in computerized tests of listening and its implications for design decisions. *Call-EJ*, *21*(1), 97–118. http://callej.org/journal/21-1/Shalmani2020.pdf

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661–699. https://doi.org/10.1111/j.0023-8333.2005.00320.x

Suryana, I., Asrianto, A., & Murwantono, D. (2020). Artificial Intelligence To Master English Listening Skills for Non-English Major Students. *Journal of Languages and Language Teaching*, *8*(1), 48. https://doi.org/10.33394/jollt.v8i1.2221

Suzuki, W., Nassaji, H., & Sato, K. (2019). The effects of feedback explicitness and type of target structure on accuracy in revision and new pieces of writing. *System*, *81*, 135–145. https://doi.org/10.1016/j.system.2018.12.017

Szarkowska, A., & Gerber-Morón, O. (2018). Viewers can keep up with fast subtitles: Evidence from eye movements. *PLoS ONE*, *13*(6). https://doi.org/10.1371/journal.pone.0199331

Tschichold, C., & Schulze, M. (2016). Intelligent CALL and written language. In F. Farr & L. Murray (Eds.), *The Routledge Handbook of Language Learning and Technology*. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315657899-44/intelligent-call-written-language-cornelia-tschichold-mathias-schulze?context=ubx&refId=5b7d3b36-7a76-47fb-9e8f-1cb5b46f5e72

von Glasersfeld, E. (1991). An Exposition of Constructivism: Why Some Like it Radical. In *Facets of Systems Science* (pp. 229–238). Springer. https://doi.org/10.1007/978-1-4899-0718-9_14

von Glasersfeld, E. (2007). *Key works in radical constructivism* (M. Larochelle (Ed.)). Sense Publishers. https://brill.com/view/title/37168?contents=toc-50344

von Glasersfeld, E., & Cobb, P. (1983). Knowledge as Environmental Fit. *Man-Environment Systems*, *13*(5), 216–224. http://vonglasersfeld.com/084

Wagner, E. (2010). The effect of the use of video texts on ESL listening testtaker performance. *Language Testing*, *27*, 493–513. https://doi.org/10.1177/026553220 9355668

Wahyuni, I., & Septianasari, L. (2020). Exploring the Efficacy of Multimedia Online-based Learning (techno-education) for Intermediate EFL Learners with Difficulties in Listening and Speaking. *Journal of Physics: Conference Series*, *1477*(4). https://doi.org/10.1088/1742-6596/1477/4/042044

Weaver, C. H. (1972). *Human listening: Processes and behavior*. The Bobbs-Merrill Company. https://eric.ed.gov/?id=ED082238

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning and Technology*, *14*(1), 65–86. http://llt.msu.edu/vol14num1/winkegasssydorenko.pdf

Wolvin, A. D. (2009). Listening, Understanding, and Misunderstanding. In W. F. Eadie (Ed.), *21st Century Communication: A Reference Handbook 21st century communication: A reference handbook* (pp. 137–146). SAGE Publications, Inc. https://doi.org/10.4135/9781412964005.n16

Wolvin, A. D. (2018). Listening Processes. *The TESOL Encyclopedia of English Language Teaching*, 1–7. https://doi.org/10.1002/9781118784235.eelt0569

Wolvin, A. D., Coakley, C. G., Wovin, A. D., Coakley, C. G., Wolvin, A. D., & Coakley, C. G. (2000). Listening Education in the 21st Century. *International Journal of Listening*, *14*(1), 143–152. https://doi.org/10.1080/10904018.2000.10499040

Worthington, D. L., & Bodie, G. D. (Eds.). (2017). *The Sourcebook of Listening Research*. John Wiley & Sons, Inc. https://doi.org/10.1002/9781119102991

Worthington, D. L., & Fitch-Hauser, M. E. (2018). *Listening* (2nd ed.). Routledge. https://doi.org/10.4324/9781315389202

Yang, Y., Yang, G., & Li, Y. (2022). The Interactive Model of L2 Listening Processing in Chinese Bilinguals: A Multiple Mediation Analysis. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.871349

Yilmaz, Y., & Granena, G. (2021). Implicitness and explicitness in cognitive abilities and corrective feedback. *Studies in Second Language Acquisition*, *43*(3), 523–550. https://doi.org/10.1017/S0272263120000601