

Audio Feedback in ESL/EFL Writing Contexts: A Review of the Literature

August 2024 – Volume 28, Number 2

<https://doi.org/10.55593/ej.28110a9>

Listiani Listiani

University of Szeged, Hungary

University of Muhammadiyah Purwokerto, Indonesia

<listianiriyanto@gmail.com>

Marianne Nikolov

University of Pécs, Hungary

<nikolov.marianne@pte.hu>

Ágnes Hódi

University of Szeged, Hungary

<hodi.agnes@szte.hu>

Abstract

Over the past two decades, multiple empirical studies have examined how teacher audio feedback works in EFL/ESL writing courses. This review explores instructor audio feedback studies in higher education writing contexts. Seventy empirical studies were identified in literature searches published between 2000 and 2022 in several academic databases. Then, based on exclusion and inclusion criteria, thirteen relevant studies were reviewed. The results indicate that interest in instructor audio feedback research has risen since 2000. This overview presents what empirical studies on teacher audio feedback in L2 writing courses aimed to find out, how they were conducted, and what they found. As many research designs and findings were underexplored, more studies are required to investigate this feedback type to benefit L2 writing learning and teaching and enrich studies on audio feedback practices. The gaps identified in the selected studies offer ideas for future explorations of teacher audio feedback in L2 writing contexts.

Keywords: L2 writing classrooms, Higher education, Teacher audio feedback

Feedback is an important part of assessment *for* learning, as it provides information on how to help learners develop (Heritage, 2019) by meeting their needs (Black et al., 2004). Feedback provides learners with important information to understand their abilities and how realistic their goals are. Feedback given by a teacher or other agents (e.g., peer, parent, self) is considered to

have the strongest impact on students' learning and achievement (Hattie & Timperley, 2007). Feedback refers to a process that impacts learning; otherwise, if it is without effect, it is only input (Dawson et al., 2018). Students not only receive feedback as information, but they also use it to improve their learning. However, feedback given after an assessment reduces its benefits to improve students' performance, as they tend to be more interested in the final score (Hughes, 2021).

Provision of feedback to learners has evolved along the rapid growth of technology. Technology advancements have encouraged teachers and researchers to implement innovations that can help students learn more efficiently. Mediated by technology, teachers can provide electronic feedback to students synchronously (e.g., online chat/online conference) or asynchronously (e.g., e-mail, comments/change tracks in Microsoft Word, recorded verbal/audio comments) (Ene & Upton, 2018; Rasi & Vuojärvi, 2018). In classroom practice, compared to oral/spoken feedback and written/text feedback, other modes, including recorded audio, audio-visual, and computer-assisted instructional feedback, are considered most effective for enhancing achievement (Hattie & Timperley, 2007, p. 84). The feedback that teachers provide has shifted from traditional ways to more modern ones (e.g., computer-mediated feedback). Prior research investigated conventional feedback provision, reporting that handwritten and oral/aural/spoken feedback delivered information to students in face-to-face settings (Elwood & Bode, 2014; Sobhani & Tayebipour, 2015). Mediated by technology, studies examined feedback in typed/electronic (e-) written, audio/verbal recorded/voice, and audio-visual feedback/screencast (Alharbi, 2022; Bakla, 2020). Audio feedback, also known as recorded oral or voice feedback, refers to a teacher's oral feedback that is recorded in e-devices (e.g., Kaizena) and delivered to students through non-digital and digital tools (e.g., tape cassette, computer, mobile phones) as an audio file or voice message. Providing teacher audio feedback to learners can be embedded in students' work (e.g., Google Docs, PDF, Word file) or sent as a separate file (e.g., e-mail, Telegram, WhatsApp).

Teacher audio feedback has been of interest to researchers to see its potential for teaching L2 learners in ESL/EFL writing classrooms. Earlier studies investigated the impact of audio feedback on learning outcomes (e.g., Rassaei, 2019; Saeed et al., 2022; Solhi & Eğinli, 2020), explored L2 learners' experiences with teacher audio feedback and their perceptions, preferences, and views about feedback (e.g., Alharbi & Alghammas, 2021; Mohammed, 2021), and examined the relationship between teacher feedback and student uptake (Alharbi, 2022; Bakla, 2020). Reviews on teacher audio feedback have been conducted to reveal its potential (Dixon, 2015) and to analyze teachers' perceptions of the innovative attributes of audio feedback, such as its observability, compatibility, complexity, trialability, and relative advantage (Killoran, 2013). These reviews contribute findings related to the strengths of audio feedback implementation that Dixon (2015) examined in tertiary education and Killoran (2013) in public education in L1 and L2 contexts. However, research designs and findings of instructor audio feedback studies in L2 writing classrooms in higher education contexts have not been studied. Therefore, those studies published between 2000 and 2022 are worth analyzing for their designs and findings.

Therefore, this paper aims to bridge this gap by investigating teacher audio feedback in ESL/EFL writing courses in higher education settings. We review research articles along four focal points: (1) research foci, (2) participants and settings, (3) research methodology and data sources, and (4) findings. The review presents findings on research designs and results to

benefit authors of future studies on audio feedback practices and instructors of L2 writing courses implementing audio feedback in higher education. Accordingly, this review aims to answer the following research questions (RQs):

RQ 1: What research questions did the studies aim to answer?

RQ 2: Who were the participants?

RQ 3: What methodology and types of data were used?

RQ 4: What were the results?

Methods

To address the aims, we reviewed empirical articles investigating teacher audio feedback in L2 writing classrooms through a literature search. We included empirical studies that gathered and analyzed new data directly, either with a qualitative or quantitative approach, including corpus studies (Riazi et al., 2018, p. 43). The search was guided by specific inclusion and exclusion criteria presented in Table 1. We included empirical studies involving university teachers, researchers, and L2 learners; they were published in English in peer-reviewed journals from 2000 to 2022. This time span allows us to understand how audio feedback research evolved.

Table 1. Inclusion Criteria in Searching Studies

Year	2000 –2022
Feedback providers	Teachers and/or researchers
Feedback receivers	ESL/EFL learners
Feedback type	Teacher audio feedback
Language	Published in English
Search database	Science Direct, Taylor and Francis Online, ERIC, Scopus
Research	Empirical
Articles	Full research articles (excluding book chapters)

To identify the relevant literature, searches were conducted in several academic databases that provide a substantial number of scholarly articles by education researchers, including Science Direct, Taylor and Francis Online, ERIC, and Scopus. The last search was conducted on 11/30/2022. We used keywords with several word combinations: audio feedback OR voice feedback OR recorded oral feedback AND English as a foreign language OR English as a second language OR ESL OR EFL OR L2 AND writing OR composition. First, 70 studies were identified that provided full texts. Then, the studies were checked for their content. We included studies on teacher audio feedback with L2 learners and university teachers or researchers in English classes where students worked on writing assignments. Fifty-seven studies were excluded, as they did not meet our criteria. They focused on written works in other fields of discipline (N = 56) and a composition course with mixed L1 and L2 learners (N = 1). Finally, thirteen studies met all criteria (Table 1).

Table 2 lists the thirteen articles divided into three time periods: 2000 – 2010, 2011 – 2020, and 2021 – 2022. The studies explored teacher audio feedback either as the only feedback mode or as one of the feedback modes (oral/spoken, text/written, and audio-visual) in writing classes at undergraduate and graduate levels. As shown in Table 2, only six studies were published on teacher audio feedback in undergraduate L2 writing contexts between 2000 and 2020. Interest in the topic increased in 2021 and 2022: an additional seven studies were published in two years

both at undergraduate and graduate levels, a slightly higher in number than in the earlier two decades. This finding shows that audio feedback has gained popularity recently.

Table 2. Thirteen Selected Articles by Year of Publication and Authors

Years	Authors
2000 – 2010	Huang (2000a, 2000b); Morra & Asís (2009)
2011 – 2020	Rassaei (2019); Solhi & Eğinli (2020); Bakla (2020)
2021 – 2022	Mohammed (2021); Tabrizi & Ranjbaran (2021); Alharbi & Alghammas (2021); Alharbi (2021, 2022); Mujtaba et al. (2022); Saeed et al. (2022)

Coding Procedure

The selected articles were read thoroughly and coded along three themes: research focus (RQ 1), contexts and participants (RQ 2), research methodology and data sources (RQ 3). The coding for these themes in the texts was also double-checked. The findings of the thirteen articles were deduced to answer the fourth research question, and the results were explained based on the primary categories found in the research foci (RQ 1).

Research focus. The first research question concerns what studies focused on. We coded the research foci into main and sub-categories identified in the abstract, introduction, and research questions of the articles, as suggested by Riazi et al. (2018, p. 44). Further identification of the research foci categories was also conducted in the other sections of the articles, such as research aims, data collection method, and data analysis method to gain a deeper understanding. To maintain the consistency of the main categories and sub-categories of the research foci, we used an approach proposed by Braun and Clarke (2006): the theme *research focus* used in this review was driven from data or using a data-driven thematic approach. We used “the authors’ actual words to describe the research foci,” as suggested by Riazi et al. (2018, p. 44).

The articles with similar or different research interests were identified to inform us of the categories and sub-categories and to note identical research foci. Further identification was also conducted on some articles that had multiple research foci to determine whether there was a main research focus. A closer reading was also conducted to double-check the research foci. After following these stages, we identified the foci in the thirteen articles.

Contexts and participants. To answer the second research question, we collected data on the contexts and participants in the information presented in the original publications. They were identified in the abstract, introduction, and research method of the articles. We adopted the term of contexts and its sub-categories used in Riazi et al. (2018, pp. 43-44) study. They sub-categorized the contexts into micro-contexts (programs) and macro-contexts (countries) in which the studies were conducted. We coded the macro-contexts based on where the studies were implemented. However, for the micro-contexts, as we coded studies that included writing programs situated in English as a foreign language (EFL) or English as a second language (ESL) contexts. The sub-categories within the programs included English composition classes (English writing course) mentioned explicitly by the authors or general courses (English course) without writing in the name. The latter included some writing instruction, writing assignments or assessments.

Participants in this review are distinguished by *educational status* and *level*. Although the terms used to define participants in the present study are like those in Riazi et al. (2018, p. 44), the scope of participants is somewhat different. We refer to *educational status* to include

instructors, learners, or instructors and learners. Concerning levels, we reviewed articles involving participants studying at undergraduate and graduate levels; these facts were stated explicitly or implicitly in the publications. Participants were grouped as instructors, undergraduates, graduates, and university students. The term “university students” refers to participants whose educational levels were not clarified or only indicated by their age range. The levels of learners’ English proficiency were also identified and reported according to the authors’ actual words in the articles.

Research methodology and data sources. The research methodology and data sources in the 13 articles were reviewed to answer the third research question. They were coded for the research design, data collection, and data analysis by adopting the categories of the research methodology established by Riazi et al. (2018, p. 45) and the classifications proposed by Hyland (2016, pp. 117-119) for data sources, as they were defined clearly. The guidelines helped us code the research methodology orientation: the codes and definitions refer to qualitative, quantitative, eclectic (QUAL + quan, QUAN + qual, QUAL + QUAN), and mixed methods (Table 3).

Table 3. Codes of Research Methodology

Code	Definition
Qualitative	Studies were purely qualitative in terms of data collection and analysis.
Quantitative	Studies were purely quantitative in terms of data collection and analysis.
Eclectic	Studies used a combination of qualitative and quantitative data and analysis but did not explicitly mention that their study was mixed methods and did not draw on the literature of mixed methods research (MMR) to frame their study. These publications were coded using one of the following combinations.
Eclectic (QUAL + quan)	The research followed a predominantly qualitative methodology but had a small portion of quantitative data and analysis (e.g., frequencies or percentages of categories).
Eclectic (QUAN + qual)	The research followed a predominantly quantitative methodology but included a small set of qualitative data and analysis (e.g., interview data).
Eclectic (QUAL + QUAN)	The research gave similar weight to both qualitative and quantitative data and analysis.
Mixed methods	The research explicitly stated that it used a mixed methodology and drew on the relevant literature to frame the study and define its purpose.

Table 4 includes the classifications and scopes of data sources. We used the term “data sources,” to refer to “ways of collecting data” following Riazi et al. (2018, p. 45) who adopted Hyland’s (2016, p. 117) classifications of *elicitations, introspection, observation, and text samples*. These different data sources were defined comprehensively by Hyland (2016, pp. 117-119). We identified how the studies collected data and coded them based on these definitions. If a combination of multiple data sources from different categories was found, we coded them according to the classifications to show that data were collected from different data source categories.

Table 4. Classifications and Scopes of Data Sources

Classification	Scopes
Elicitation	Study included self-reports and performance data such as questionnaires, interviews, focus groups, and tests.
Introspection	Study analyzed verbal or written reports including think-aloud protocols, retrospective reports, and diaries.
Observation	Study comprised directed or recorded data of live interactions or writing behaviours including audio or video recording or keystroke logging.
Text samples	Study included a sample of naturally produced samples of writing including single or chains of texts or corpora.

Results and Discussion

This section presents the review results of thirteen selected articles investigating teacher audio feedback in ESL/EFL writing courses at undergraduate and graduate levels. We set four aims: to analyze the research foci, the participants and the settings in which the studies were conducted, the research methodology and data sources researchers used, and the results researchers found in their studies. These are presented in the next sections.

Research Foci

Table 5 lists the primary and sub-categories of research foci in the thirteen papers. After coding the research foci of the studies, we identified five primary and twelve sub-categories of research foci that they aimed to examine. The thirteen studies were classified into five groups of main categories. Therefore, the number of studies of each main category was based on the total of studies in which their research aims were classified into.

Table 5. Research Foci in 13 Studies

Authors	Primary category	Sub-categories	Total (%)
Saeed et al. (2022); Solhi & Eğinli (2020); Tabrizi & Ranjbaran (2021); Alharbi (2022)	Feedback effect	Writing performance, writing accuracy	3 (23.08)
Mujtaba et al. (2022); Rassaei (2019)	Feedback uptake	Text revisions	1 (7.69)
Alharbi (2021)	Feedback and perceptual style effect	Writing accuracy	2 (15.38)
Morra & Asís (2009)	Feedback effect	Writing performance, student perception	1 (7.69)
Mohammed (2021)	Feedback uptake	Error correction, student perception	1 (7.69)
Alharbi & Alghammas (2021); Huang (2000a)	Student perspective	Student preference, factors shaping preferences	1 (7.69)
Huang (2000b)	Feedback efficiency	Feedback quantity, feedback content, student perception, student preferences, reasons for preferences	2 (15.38)
Bakla (2020)	Feedback effect	Feedback quantity, feedback focus, feedback provision process	1 (7.69)
		Text revision, student preference, reasons for preferences, feedback acceptance process	1 (7.69)

The main categories include perceptual style, feedback effect, feedback uptake, student perspective, and feedback efficiency. The sub-categories include writing performance, writing accuracy, text revisions, student perception, error corrections, student preferences, factors shaping preferences, feedback quantity, feedback content, feedback focus, feedback provision process, and feedback acceptance process. The terms are self-explanatory; however, a few

terms require more elaboration. For example, *perceptual style*, according to Rassaei (2019), referred to different types of individuals (auditory, visual, and read/write) who had a specific way of absorbing and processing information, which was considered a factor that moderated feedback effect. *Feedback effect* focused on the impact of teacher feedback on students' writing performance (e.g., Solhi & Eğinli, 2020) and writing accuracy (e.g., Tabrizi & Ranjbaran, 2021). *Feedback uptake* concerned to what extent students integrated feedback in their text revisions (Alharbi, 2022) or used feedback to reduce the number of mistakes (Morra & Asís, 2009). *Feedback efficiency* measured how feedback was considered efficient for a teacher to deliver feedback (e.g., Huang, 2000a). *Feedback provision process* referred to how the nature of feedback delivery among feedback modes differed (Huang, 2000b). *Feedback acceptance process* emphasized the process involved in accepting and acting upon feedback.

We found that research on *the effects of instructor feedback modes on learner writing and learner perspectives towards feedback modes* were the most frequently studied by five and six studies, respectively (Table 5). The studies focused only on measuring feedback effects on student writing (e.g., Solhi & Eğinli, 2020) combined it with the effect of student perceptual style to investigate how feedback impacted student writing (Mujtaba et al., 2022; Rassaei, 2019), or included learner perspectives to support the effects (Alharbi, 2021; Bakla, 2020). Only one study (Mohammed, 2021; 7.69%) focused on learner perspectives, but the same focus was also investigated in studies exploring feedback efficiency (Alharbi & Alghammas, 2021; Huang, 2000a) and feedback uptake (Morra & Asís, 2009). However, feedback uptake (Alharbi, 2022) and feedback efficiency (Huang, 2000b) alone were less frequently studied. Though *feedback provision process* and *feedback acceptance process* were included in two sub-categories and research focus shifted to *feedback uptake*, *feedback efficiency*, and *perceptual style effect*, the focus of *feedback effect* and *student perspectives* were popular in studies for over two decades. Given the narrow focus of earlier studies, more research is needed to investigate complex phenomena related to teacher audio feedback practices in depth.

Contexts and Participants

Findings on contexts (countries and programs) and participants (educational status and levels) in which the studies were conducted are presented in Tables 6, 7, and 8. Table 6 summarizes the countries (macro-contexts) where the studies explored instructor audio feedback in ESL/EFL writing courses across three intervals. Data show the changes in the number of articles (3, 3, 7) and the periods in years (11, 10, 2) using frequency counts and percentages. We found that Arab Saudi was dominant, accounting for 38.46% (n = 5) in the third time interval. It was followed by Taiwan, Iran, and Turkey, two studies were published on them (15.38%), whereas one study (7.69%) was conducted in Argentina and Pakistan, respectively.

Table 6. Countries Where Research was Conducted

Authors	Macro contexts (countries)	2000-2010	2011-2020	2021-2022	Total (%)
Huang (2000a, 2000b)	Taiwan	2 (15.38)	0 (0)	0 (0)	2 (15.38)
Morra & Asís (2009)	Argentina	1 (7.69)	0 (0)	0 (0)	1 (7.69)
Rassaei (2019); Tabrizi e,g,(2021)	Iran	0 (0)	1 (7.69)	1 (7.69)	2 (15.38)
Alharbi. (2021, 2022); Alharbi & Alghammas (2021); Mohammed (2021); Saeed et al. (2022)	Saudi Arabia	0 (0)	0 (0)	5 (38.46)	5 (38.46)
Bakla (2020); Solhi & Eğinli (2020)	Turkey	0 (0)	2 (15.38)	0 (0)	2 (15.38)
Muhammad Mujtaba et al. (2022)	Pakistan	0 (0)	0 (0)	1 (7.69)	1 (7.69)
Total		3 (23.07)	3 (23.07)	7 (53.84)	13 (100)

Table 7 illustrates the writing programs where teacher audio feedback was employed. The feedback mode was studied predominantly in EFL contexts: twelve publications were (92.30%) in the database, including nine studies in writing courses and three studies in English courses. Only one paper (7.69%) gave an account of a study published in 2022 in an ESL English course. It is interesting to note that the number of articles focusing on EFL programs was stable over the first two time periods (3, 23.08% and 3, 23.07%, respectively), but there was an increase in articles in both EFL and ESL contexts in the last two years (7, 53.84%). Therefore, L2 writing researchers have worked with teacher audio feedback more often in EFL courses than in ESL contexts.

Table 7. Programs Where Research was Conducted

Authors	Micro contexts (Programs)	2000-2010	2011-2020	2021-2022	Total (%)
Alharbi (2021, 2022); Bakla (2020); Huang (2000a, 2000b); Mohammed (2021); Morra & Asís (2009); Saeed et al. (2022); Solhi & Eğinli (2020)	EFL Writing course	3 (23.08)	2 (15.38)	4 (30.77)	9 (69.23)
Alharbi & Alghammas (2021); Rassaei (2019); Tabrizi (2021)	English course	0 (0)	1 (7.69)	2 (15.38)	3 (23.07)
Mujtaba et al. (2022)	ESL English course	0 (0)	0 (0)	1 (7.69)	1 (7.69)
Total		3 (23.08)	3 (23.07)	7 (53.84)	13 (100)

Table 8 shows the educational status and levels of participants in 13 studies. Most participants were students, but instructors were also included as participants in two publications: both were conducted by Huang (2000a; 2000b). Nine studies (69.22%) specifically mentioned undergraduate as their participants' educational level (e.g., Alharbi, 2021), whereas the other three studies (23.07%) stated the ranges of their age and were categorized as studying at the undergraduate level (i.e., Rassaei, 2019; Solhi & Eğinli, 2020; Tabrizi, 2021). A study conducted by Mohammed (2021) specifically mentioned graduate level. While examining *undergraduates* was the focus in ten studies in the three periods (1, 7.69%; 3, 23.07%; and 6, 46.15%, respectively), there was an interest in *graduate* participants in the most recent years. Less interest was found in studying *instructors and undergraduates* over the most recent time periods (2, 15.38%; 0, 0%; and 0, 0%).

Among 13 studies, six studies failed to mention their participants' English proficiency specifically. Two studies on undergraduates (Alharbi, 2022; Bakla, 2020) and university students (Rassaei, 2019; Tabrizi, 2021) classified the students' level of English proficiency as intermediate. Two studies that involved undergraduates (Morra & Asís, 2009; Mujtaba et al., 2022) and one study on unspecified level university students (Solhi & Eğinli, 2020) classified their participants at the upper intermediate level. However, six studies did not mention what level the students were at, including three studies on undergraduates (Alharbi, 2021; Alharbi & Alghammas, 2021; Saeed et al., 2022), a study conducted at the graduate level (Mohammed, 2021), and two studies that involved undergraduates and their instructor (Huang, 2000a; 2000b). Most studies considered their participants' level of English proficiency as a key variable when examining instructor audio feedback, despite it being slightly higher in number than those without considering participants' proficiency level. Thus, they focused on students at the intermediate and upper intermediate levels, whereas no publication involved learners of English at the beginner and advanced proficiency levels.

Table 8. Educational Status and Levels of Participants

Authors	Participants (Educational status & levels)		2000-2010	2011-2020	2021-2022	Total (%)
Alharbi (2021, 2022); Alharbi & Alghammas (2021); Bakla (2020); Morra & Asís (2009); Mujtaba et al. (2022); Saeed et al. (2022)	Students	Undergraduates	1 (7.69)	1 (7.69)	5 (38.46)	7 (53.84)
Mohammed (2021)	Students	Graduates	0 (0)	0 (0)	1 (7.69)	1 (7.69)
Rassaei (2019); Solhi & Eğimli (2020); Tabrizi (2021)		University students	0 (0)	2 (15.38)	1 (7.69)	3 (23.07)
Huang (2000a; 2000b)	Instructor, students	Instructor, undergraduates	2 (15.38)	0 (0)	0 (0)	2 (15.38)
Total			3 (23.07)	3 (23.07)	7 (53.84)	13 (100)

In sum, typical research contexts in studies investigating teacher audio feedback in EFL/ESL writing contexts between 2000 and 2022 were overwhelmingly undergraduate programs in EFL contexts. They mostly focused on analyzing data from the students' perspectives. The fact that studies investigated students rather than teachers, or both in interaction, shows that there was little interest in involving university instructors and how they use audio feedback.

Research Methodology and Data Sources

Table 9 summarizes the orientations of research methods in the 13 studies. The most frequently used methodology was a quantitative approach (e.g., Alharbi, 2021), counting six (46.15%) studies. It was followed by four (30.76%) inquiries that combined quantitative and qualitative methods with one prominent method or with the same weight given to both methods, such as in studies conducted by Huang (2000a, 2000b) and Morra and Asís (2009). The least frequent approaches were mixed methods (e.g., Bakla, 2020) and qualitative research (Alharbi & Alghammas, 2021), 2 (15.38%) and 1 (7.69%), respectively. Based on these findings, we can state that most projects used quantitative or a combination of quantitative and qualitative methods to answer their research questions. As for their frequencies over the years, the latter approach decreased in the number of articles in over the three periods (3, 23.07%; 0, 0%; and 1, 7.69% respectively), whereas mixed and qualitative methods increased in the last two-time intervals (0, 0%; 1, 7.69%; 2, 15.38%, respectively).

Table 9. Methodological Orientations

Authors	Research methodology orientations	2000-2010	2011-2020	2021-2022	Total (%)
Alharbi (2022); Mujtaba et al. (2022); Rassaei (2019); Solhi & Eğimli (2020); Tabrizi (2021); Saeed et al. (2022);	Quantitative	0 (0)	2 (15.38)	4 (30.77)	6 (46.15)
Alharbi & Alghammas (2021)	Qualitative	0 (0)	0 (0)	1 (7.69)	1 (7.69)
Bakla (2020); Mohammed (2021)	Mixed	0 (0)	1 (7.69)	1 (7.69)	2 (15.38)
Huang (2000b); Morra & Asís (2009)	Eclectic (QUAN+QUAL)	2 (15.38)	0 (0)	0 (0)	2 (15.38)
Alharbi (2021); Huang (2000a)	Eclectic (QUAN+qual)	1 (7.69)	0 (0)	1 (7.69)	2 (15.38)
Total		3 (23.07)	3 (23.07)	7 (53.84)	13 (100)

The most frequent data source used in the published studies was elicitations, as shown in Table 10. They were used in seven (53.84%) studies (e.g., Alharbi, 2021; Mohammed, 2021; Rassaei, 2019) and an increase was found in the number of articles in the last two periods (0, 0%; 2, 15.38%; and 5, 38.46%, respectively). A combination of elicitations and different data sources, such as text samples, was found in four (30.77%) studies (e.g., Alharbi & Alghammas, 2021) and observation in one (7.69%) study (Bakla, 2020). Text samples alone were used in one study (7.69%). Though elicitations were the most dominant data source over the years, text samples were combined with them in 2022. This trend indicates that authors considered authentic samples of students' written texts important to be integrated in their studies.

Table 10. Data Source Classifications Used

Authors	Data sources	2000-2010	2011-2020	2021-2022	Total (%)
Alharbi (2021); Mohammed (2021); Mujtaba et al. (2022); Rassaei (2019); Saeed et al. (2022); Solhi & Eginli (2020); Tabrizi (2021)	Elicitations	0 (0)	2 (15.38)	5 (38.46)	7 (53.84)
Alharbi (2022)	Text samples	0 (0)	0 (0)	1 (7.69)	1 (7.69)
Alharbi & Alghammas (2021); Huang (2000a, 2000b); Morra & Asís (2009)	Text samples, elicitation	3 (23.08)	0 (0)	1 (7.69)	4 (30.77)
Bakla (2020)	Elicitations, observation	0 (0)	1 (7.69)	0 (0)	1 (7.69)
Total		3 (23.08)	3 (23.07)	7 (53.84)	13 (100)

The data sources presented in Table 10 were further analyzed in their categories to indicate the scopes of data sources; they are presented in Table 11. The 13 studies predominantly used a combination of data source categories that involved either a questionnaire or interview, or both. They were integrated with one or more categories from different data sources, such as written drafts (e.g., Morra & Asís, 2009) and screen recordings (Bakla, 2020), or from within similar data sources, such as writing tasks/tests (e.g., Alharbi, 2021). A combination of *questionnaire* and *interview* was used in one (7.69%) study (Mohammed, 2021). Three (23.08%) studies (e.g., Saeed et al., 2022) used only *writing tests*, and one (7.69%) study (Alharbi, 2022) analyzed *written drafts*.

Table 11. The Scopes of Data Source Used

Authors	Scopes of data sources	Total (%)
Huang (2000a, 2000b); Morra & Asís (2009)	Writing draft, questionnaire	3 (23.08)
Alharbi & Alghammas (2021)	Writing draft, interviews	1 (7.69)
Mohammed (2021)	Questionnaire, interview	1 (7.69)
Alharbi (2021); Mujtaba et al. (2022); Rassaei (2019)	Questionnaire, writing test/task	3 (23.08)
Saeed et al. (2022); Solhi & Eginli (2020); Tabrizi (2021)	Writing test	3 (23.08)
Alharbi (2022)	Writing draft	1 (7.69)
Bakla (2020)	Questionnaire, interview, writing task, screen recordings	1 (7.69)

In summary, teacher audio feedback studies in L2 writing contexts predominantly used quantitative research design. Combining quantitative and qualitative research was the second frequent method applied in the publications. The use of multiple categories, either from a similar or different data source, was also their preference, as authors wanted to provide valid, reliable, and authentic pictures of complex situations, implementing what Hyland (2016, p. 121)

said, “the use of multiple sources of data or analytical methods, can bring greater plausibility to the interpretation of results.” We found that qualitative research was rarely used, although a tendency to shift towards using both quantitative and qualitative datasets was clear in teacher audio feedback studies published over the past two decades.

Main Findings in Publications

The findings of the 13 studies are presented in this section: they documented varied outcomes. They are explained based on the primary category of their research focus.

Feedback effect. The effect of feedback was the focus of the audio feedback inquiries. They employed *writing tests* or *writing tasks* and evaluated them to determine how feedback impacted the outcomes. Three studies investigated the effect of teacher audio feedback and written feedback on learners' writing performance and accuracy. Solhi and Eğinli (2020) found that the group receiving teacher audio feedback outperformed, especially in content and organization, the group that received metalinguistic written corrective feedback. However, in clarity and sentence-level accuracy, neither of the groups showed significant differences. The findings were supported by Tabrizi's (2021) study. It showed that students receiving computer-mediated (CM) audio-based corrective feedback (CF) also outperformed those receiving CM text-based CF in terms of writing accuracy. Alharbi (2021), who investigated learners' writing performance, found that students who got teacher audio feedback outperformed their peers who got teacher-written feedback in writing argumentative texts.

The effect of audio feedback and other feedback modes, such as written/text, spoken, screencast, or audio-visual, was investigated in two studies. Bakla (2020) showed that the effect of teacher audio feedback resulted in the highest number of correct revisions in the essay writing tasks when students got teacher written feedback and screencast. In a study conducted by Saeed et al. (2022), students who received teacher audio feedback, spoken feedback, or audio-visual feedback outperformed a group receiving teacher written feedback on their performance of paragraph writing.

Mediated by students' preferred perceptual style, the effectiveness of CM feedback was promoted when it was aligned with their preferred perceptual style. Both CM audio feedback and CM written feedback showed effective for developing learners' L2 and writing accuracy. However, CM audio-based CF was more effective than CM text-based CF for students' L2 development, particularly in using English articles (Rassaei, 2019), and the CM audio feedback also proved to be more effective for students' writing accuracy in using past perfect tense than CM text feedback (Mujtaba et al., 2022).

In conclusion, teacher audio feedback exerted more positive effects than teacher written feedback and screencast in improving learners' writing performance, accuracy, and the number of correct revisions, but its impact was like spoken and audio-visual feedback modes in enhancing writing performance and accuracy. In addition, students' preferred perceptual styles could promote the effectiveness of CM feedback, but CM audio-based feedback showed more positive results than CM text-based feedback.

Student perspectives. Either questionnaires and interviews or one of the two data sources were used to collect data on students' perspectives in earlier studies, and the findings indicated varied outcomes. Some studies found that learners did not show preference, and perceived teacher audio feedback more positively, but the studies only highlighted the benefits and drawbacks of the feedback modes they investigated. For example, the study conducted by Bakla (2020) found

that students did not uniformly prefer a particular feedback mode from among three feedback modes. Students' priorities, such as practicality, comprehensibility, multimodality, effectiveness, interactivity, and the researcher's social presence influenced their preference for a particular feedback mode. However, their preference did not relate to a single factor of benefits alone; one factor could result from several combined factors. In Mohammed's (2021) study, the students' preferences for feedback modes (oral, e-written, audio, and screencast) varied according to affordance and limitations. Mohammed also found some factors that shaped the learners' preferences that referred to comprehensibility, multimodality, interactivity, specificity of feedback, revision settings, devices to access feedback, internet connection, learners' knowledge of genre and errors, previous experience, and individual differences in learning styles. Morra and Asís (2009) found that their respondents perceived both teacher audio feedback and written feedback as helpful in revising their papers, particularly feedback focusing on micro errors (vocabulary, grammar, mechanics) was more beneficial than pointing out macro errors (content, organization).

Compared to other feedback modes, some studies showed that teacher audio feedback was perceived as the most beneficial for some reasons, such as its personalization and details. On the other hand, teacher written feedback was preferred over audio feedback because of some factors, such as clarity and accessibility. Alharbi (2021) showed that learners perceived audio feedback as more efficient in details, clarifications, and personalization than written feedback that was perceived as clearer and easier to understand and interpret. Overall, both feedback modes were found to be equally satisfactory and accessible. Furthermore, Huang (2000a) found that learners perceived more positively and preferred teacher audio-taped feedback (ATF) over written feedback only (WF-only). In contrast with Alharbi (2021) and Huang (2000a), Alharbi and Alghammas (2021) indicated that teacher written feedback received higher preference over teacher audio feedback for its clarity, easiness, easy access to feedback, and its focus on a particular issue in the assignments. Some challenges were also highlighted by the students regarding feedback length, detailed instruction, and the difficulty in accessing feedback.

In summary, learners' perspectives on feedback modes investigated in six studies emphasized their perception, preference, and the benefits and downsides of feedback modes. Three studies (Bakla, 2020; Mohammed, 2021; Morra & Asís, 2009) showed that their respondents did not uniformly prefer a particular feedback mode or perceive a specific mode more positively than another mode. However, the learners only highlighted the strengths and challenges of the feedback modes they received. Alharbi (2021) and Huang (2000a) indicated their respondents' preference and positive perceptions to teacher audio feedback over written feedback with some reasons shaping their preference; preference of teacher written feedback over audio feedback was also found, but the number of students selecting each feedback mode was only slightly different (Alharbi & Alghammas, 2021).

Feedback efficiency. Earlier studies measured how efficient the feedback modes were in terms of the quantity of feedback and content collected from learners' writing. Two studies quantified the number of words in the feedback and the time used to deliver feedback. Huang (2000a; 2000b) investigated a combined method consisting of written feedback and ATF, and WF-only. Huang found that the former feedback mode and ATF-only were much more efficient than the WF-only. In contrast to the findings by Alharbi and Alghammas (2021), though they also counted the number of audio feedback and written feedback, the number of comments provided in audio feedback was lower than in written feedback. However, their investigation related to

whether each comment of both feedback modes fulfilled one or more language functions. Results showed that the number of language functions carried out by each comment in audio feedback mode was higher than each comment in written feedback mode. Audio feedback mode could carry from one to four different pragmatic functions, such as statement, asking question, advice, suggestion, and justification for each comment. However, written feedback mode could only carry the combination of two functions in each comment, such as suggestion and question, statement and question, praise and question, evaluation and question.

In sum, in terms of the feedback efficiency of both audio feedback and written feedback, earlier studies described a diverse picture. The quantity of feedback in terms of the number of words included in audio feedback mode was higher than that in written feedback mode, but it was lower in terms of the number of comments. However, the number of pragmatic functions of each comment provided in audio feedback mode was twice as high as the number of pragmatic functions mentioned in written feedback mode.

Feedback uptake. As for feedback uptake, L2 researchers investigated how far feedback got integrated into students' text revisions and to what extent and in what ways feedback helped them reduce errors in their final texts. Alharbi (2022) studied how far students used feedback in their text revisions and how different their integration across the feedback modes (oral/spoken, text, recorded audio, and audio-visual) was. Students implemented most (83.52%) of feedback integration, its quality reaching 68.46%. The findings on feedback quantity indicated that most students integrated feedback into their revisions from audio-visual feedback, followed by oral, audio, and the least successful one was written feedback. Learners integrated questions, suggestions, and imperatives more than other feedback features, such as correction and statement. Furthermore, Morra and Asís (2009), who investigated two feedback modes (audio and written), found a significant decrease in the number of errors and weaknesses at micro or macro levels in students' final texts regardless of feedback mode. Therefore, audio feedback could help learners either with revisions or error and weakness reductions.

Process of feedback provision and acceptance. How feedback was provided and received was also studied in the chosen publications. Huang (2000b) investigated the difference in provision between the combined method and written feedback collected from learners' texts. The author found a distinction of writing aspects addressed between the two feedback modes. When a combined method was used, written feedback commented on language errors, and ATF was applied to address both language errors and other problems concerning content, structure, organization, coherence, logic, clarity, tone, and style. Furthermore, if only ATF was used, it also commented on language errors and the aspects unrelated to language errors more thoroughly than only written feedback. In addition, ATF also encouraged the use of L1 to explain writing problems, whereas written feedback did not. ATF provided information on more pragmatic functions in comments than written feedback, including statement, question, suggestion, and advice. Unlike Huang (2000b), Bakla (2020) investigated the different processes of receiving feedback among three digital feedback modes (written, audio, and screencast) by analyzing screen recordings and interviews on ways of students' engagement and interaction with the three feedback modes and found that learners usually followed similar procedures and worked on similar problems. They usually used mobile phones to check the availability of teacher feedback and worked more with written feedback than audio or audio-visual. However, a computer was often chosen, as it provided a larger screen than mobile

devices. In addition, the interaction between teacher and student was poor in audio and written feedback, and there was almost no interaction when they used screencasts.

In conclusion, in terms of feedback provision, combined feedback and ATF gave feedback more thoroughly than WF, and in terms of feedback acceptance, students usually responded to audio feedback and written feedback, but they did not interact with their teacher as a source of feedback in the cases when they received screencast or audio-visual feedback.

Conclusion

The low number of selected studies indicated that instructor audio feedback in ESL/EFL writing contexts in higher education was not a popular topic in L2 writing research. Although the number of publications on audio feedback is relatively low, it increased between 2021 and 2022. Along with the research questions, the findings indicated that most studies focused on the effect of feedback and learners' perspectives of feedback, and other foci related to feedback efficiency, uptake, and provision and acceptance process were of less interest. The 13 studies were conducted in six countries in EFL and ESL undergraduate and graduate programs teaching students at intermediate and upper intermediate proficiency levels. Lower and advanced English proficiency levels are yet to be investigated.

The studies mostly collected quantitative data sometimes in combination with qualitative data. Although the findings suggest benefits to L2 learners' writing skills in specific circumstances in general, they were limited to the discussion of enhancing learners' writing performance and accuracy, text revisions and error reductions, learners' perspectives, and quantity and quality of comments. Further research is required to develop a more generalizable understanding of the subject matter, with a particular focus on the contribution of such findings to the advancement of writing pedagogy and research. Therefore, future research designs and topics of teacher audio feedback in L2 writing courses are still wide open, and using qualitative and mixed methods in different contexts on larger samples could contribute to ESL/EFL writing pedagogy in important ways.

Although this literature review suggests some important information for future education research and L2 writing pedagogy, there are some limitations that need to be addressed in future studies. The first limitation is that the studies were carried out only in higher education contexts. This implies that the findings were based on research in specific contexts with adults. Therefore, it is necessary to conduct similar reviews at lower levels of education to compare results. The second limitation of this review is that we focused only on teachers providing feedback. Studies involving peers as feedback providers can elicit more detailed information about students' views and experiences with audio feedback practices.

As for specific future studies, we would like to highlight three directions for L2 writing researchers planning to conduct research on teacher audio feedback in EFL or ESL writing contexts. First, we must look back at the typical research contexts and the participants as they were represented in this paper. As teacher audio feedback in L2 writing research was limited in number, L2 writing researchers should find more insights in these EFL or ESL contexts particularly. Conducting research in other countries, in diverse ESL/EFL writing contexts of English courses, and at different levels of learners' English proficiency can contribute to a more nuanced picture of L2 writing research. Moreover, investigating and involving instructors as feedback givers could also provide new perspectives and enrich the data. Therefore, the

contexts and participants of audio feedback studies still need to go beyond the programs where they have so far been studied.

Second, studies should shift to more complex foci to offer a more authentic picture of audio feedback practices. L2 writing researchers should explore instructors' cognition and motivation to understand what they know, think, believe, and why they act the way they do in their classroom practices (Borg, 2003). This approach would allow future research to offer insights into the ways instructors provide feedback. Researchers should also investigate in what ways and to what extent students take up their instructor's feedback for their revision and why.

Third, from the results of our analysis related to research methodology and data sources, combining qualitative and quantitative methods and using multiple data sources are important. However, we highlight that mixed methods and multiple types and sources of data could strengthen findings by using triangulation. Using qualitative methods would allow researchers to explore and describe the complex phenomenon of feedback practice, as there were hardly any such publications. These future directions can offer valuable insights and strengthen the innovative practices of audio feedback in L2 writing contexts.

About the Authors

Listiani Listiani (corresponding author) is a Ph.D. student of the Doctoral School of Educational Sciences at the University of Szeged, Hungary. Her main research interests include L2 English writing, language assessment, English learning and instruction, and computer-assisted language learning. ORCID ID: 0009-0003-2287-2694

Marianne Nikolov is a Professor Emerita of English Applied Linguistics at the University of Pécs, Hungary. Her research topics are teaching English to young learners, assessment in language education, individual differences, teachers' beliefs and practices, and language policy. Her full CV could be seen in her website: http://ies.btk.pte.hu/content/nikolov_marianne ORCID ID: 0000-0001-5844-4729

Ágnes Hódi is a literacy researcher based at the Juhász Gyula Faculty of Education, University of Szeged, Hungary. Her research focuses on issues related to literacy development, particularly in the context of early childhood education and educational technology. ORCID ID: 0000-0003-0325-1449

To Cite this Article

Listiani, L., Nikolov, M. & Hódi, Á. (2024). Audio feedback in ESL/EFL writing contexts: A review of the literature. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, 28(2). <https://doi.org/10.55593/ej.28110a9>

References

Alharbi, M. A. (2021). Impact of teacher written vs. audio feedback on EFL undergraduates' writing. *Cypriot Journal of Educational Sciences*, 16(3), 1141–1153. <https://doi.org/10.18844/CJES.V16I3.5836>

Alharbi, M. A. (2022). Exploring the impact of teacher feedback modes and features on students' text revisions in writing. *Assessing Writing*, 52(February), 100610. <https://doi.org/10.1016/j.asw.2022.100610>

Alharbi, M. A., & Alghammas, A. (2021). Teacher written vs. audio feedback on

- undergraduates' written assignments. *Theory and Practice in Language Studies*, 11(12), 1562-1570. <https://doi.org/10.17507/tpls.1112.08>
- Bakla, A. (2020). A mixed-methods study of feedback modes in EFL writing. *Language Learning and Technology*, 24(1), 107–128. <https://doi.org/10125/44712>
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8–21. <https://doi.org/10.1177/003172170408600105>
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81-109. <https://doi.org/10.1017/S0261444803001903>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Dawson, P., Henderson, M., Ryan, T., Mahony, P., Boud, D., Phillips, M. & Molloy, E. (2018). Technology and feedback design. In J. M. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, design, and technology: An International compendium of theory, research, practice and policy*. Springer. https://doi.org/10.1007/978-3-319-17727-4_124-1
- Dixon, S. (2015). The pastoral potential of audio feedback: a review of the literature. *Pastoral Care in Education*, 33(2), 96–104. <https://doi.org/10.1080/02643944.2015.1035317>
- Elwood, J. A., & Bode, J. (2014). Student preferences vis-à-vis teacher feedback in university EFL writing classes in Japan. *System*, 42, 333-343. <https://doi.org/10.1016/j.system.2013.12.023>
- Ene, E., & Upton, T. A. (2018). Synchronous and asynchronous teacher electronic feedback and learner uptake in ESL composition. *Journal of Second Language Writing*, 41, 1-13. <https://doi.org/10.1016/j.jslw.2018.05.005>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Heritage, M. (2019). Feedback for Enhanced English Language Learning. In: X. Gao (Eds.), *Second handbook of English language teaching*. Springer. https://doi.org/10.1007/978-3-030-02899-2_27
- Hyland, K. (2016). Methods and methodologies in second language writing research. *System*, 59, 116-125. <https://doi.org/10.1016/j.system.2016.05.002>
- Huang, S. Y. (2000a). A quantitative analysis of audiotaped and written feedback produced for students' writing and students' perceptions of the two feedback methods. *Tunghai Journal*, 41, 199–232. <https://files.eric.ed.gov/fulltext/ED448604.pdf>
- Huang, S. Y. (2000b). The nature of an EFL teacher's audiotaped and written feedback on student writing: A case study. (Project No: NSC 88-2411-H-029-008) [Grant]. The National Science Council of Taiwan. <https://files.eric.ed.gov/fulltext/ED438728.pdf>
- Hughes, L. (2021). Washback and the assessment practices of ESL instructors at Japanese university. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, 5(1), 1-9. <https://doi.org/10.30743/ll.v5i1.3238>

- Killoran, J. B. (2013). Reel-to-reel tapes, cassettes, and digital audio media: Reverberations from a half-century of recorded-audio response to student writing. *Computers and Composition*, 30(1), 37–49. <https://doi.org/10.1016/j.compcom.2013.01.001>
- Mohammed, M. A. S. (2021). Does teacher feedback mode matter for language students? *Asian EFL Journal*, 28(11), 202–227. <https://www.asian-efl-journal.com/wp-content/uploads/AEJ-Volume-28-Issue-1.1-February-2021.pdf>
- Morra, A. M., & Asís, M. I. (2009). The effect of audio and written teacher responses on EFL student revision. *Journal of College Reading and Learning*, 39(2), 68–81. <https://doi.org/10.1080/10790195.2009.10850319>
- Mujtaba, M. S., Parkash, R., Kaur Mehar Singh, M., & Kamyabi Gol, A. (2022). The Effect of computer-mediated feedback on L2 accuracy. Does the difference in learners' perceptual style moderate the effectiveness of the feedback? *Computers in the Schools*, 39(2), 99-119. <https://doi.org/10.1080/07380569.2022.2041891>
- Rasi, P., & Vuojärvi, H. (2018). Toward personal and emotional connectivity in mobile higher education through asynchronous formative audio feedback. *British Journal of Educational Technology*, 49(2), 292–304. <https://doi.org/10.1111/bjet.12587>
- Rassaei, E. (2019). Computer-mediated text-based and audio-based corrective feedback, perceptual style and L2 development. *System*, 82, 97–110. <https://doi.org/10.1016/j.system.2019.03.004>
- Riazi, M., Shi, L., & Haggerty, J. (2018). Analysis of the empirical research in the journal of second language writing at its 25th year (1992–2016). *Journal of Second Language Writing*, 41, 41-54. <https://doi.org/10.1016/j.jslw.2018.07.002>
- Saeed, M. A., AbuSa'aleek, A. O., & RahmtAllah, E. A. E. (2023). The impact of feedback modes on learners' performance in paragraph writing. In: K. Arai (Ed.), *Proceedings of the Future Technologies Conference (FTC) 2022*, Vol. 3. FTC 2022 2022. Springer. https://doi.org/10.1007/978-3-031-18344-7_55
- Sobhani, M., & Tayebipour, F. (2015). The effects of oral vs. written corrective feedback on Iranian EFL learners' essay writing. *Theory and Practice in Language Studies*, 5(8), 1601. <https://doi.org/10.17507/tpls.0508.09>
- Solhi, M., & Eğinli, I. (2020). The effect of recorded oral feedback on EFL learners' writing. *Journal of Language and Linguistic Studies*, 16(1), 1–13. <https://doi.org/10.17263/jlls.712628>
- Tabrizi, A. N., & Ranjbaran, M. M. (2021). The effect of computer-mediated text-based and audio-based corrective feedback on the development of writing accuracy of Iranian EFL learners. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(13), 4675-4692. <https://www.turcomat.org/index.php/turkbilmat/article/view/9639/7348>

Copyright of articles rests with the authors. Please cite TESL-EJ appropriately.