Defining Grammatical Difficulty to Make Better Choices about Corrective Feedback: A Meta-Analysis of Persian EFL Learners

Andrew Schenck1

Abstract
Past experimental studies of corrective feedback (CF) have isolated factors like grammatical complexity, learner proficiency, and L1 one by one, carefully designing experiments that eliminate the influence of “extraneous” factors. Because each factor is actually codependent, more holistic study is needed. Eleven studies, all of which had English as a Foreign Language (EFL) learners with a Persian L1 and productive measures of speech or writing, were selected for meta-analytic examination. Results suggest that type of grammatical feature, as well as associated learner variables such as L1 similarity or proficiency, collectively influence the efficacy of different CF types. As variables jointly add to the difficulty of a grammatical feature, CF providing a kind of scaffold, in the form of a written or oral reformulation from the teacher, appears to be the most effective. As grammatical difficulty decreases, learners appear to benefit from CF in which the learner is compelled to self-repair.

Keywords
corrective feedback, recasts, direct feedback, written, oral, Persian, grammar, EFL learners

Introduction
Concerning written Corrective Feedback (CF), Lee (2019a) describes an English teacher in Hong Kong who is fed up with never-ending corrections that don’t improve students’ editing skills, “as students continue to make mistakes, even the very same mistakes” (p. 524). Nearly every EFL instructor can sympathize with this dilemma, slaving over corrections, only to realize that their work has little to no effect. This phenomenon has led some researchers to suggest that correcting written errors is not effective at all (Truscott, 1996, 1999). While a great deal of research has emerged to suggest that written forms of CF are indeed effective (Bitchener et al., 2005; Chen et al., 2016; Ferris, 2006; Shintani & Ellis, 2013), controversy on how to use this pedagogical technique continues to persist. Lee (2019a) contends that problems with providing written feedback lie in a tendency to correct all errors, which does not allow for focus on specific grammatical features. Although heightened emphasis of just one target feature appears to improve student performance in writing, how a feature should be emphasized is still an issue of contention (Mao & Lee, 2020). Some studies indicate that correcting an error in writing, referred to as direct corrective feedback (DCF), leads to higher gains when used along with DCF (Bitchener et al., 2005; Sheen, 2007). Lee (2019b) accurately points out that such discrepancies are the result of heavy reliance on experimental studies, which do not result in “research that can guide teachers’ actual practice” (p. 1). Experimental research conducted so far does limit the adaptation of theory to practice, explaining why more qualitative research has been recommended to help bridge this gap.

Although it is true that more qualitative research is needed, reliance on experimental study does not embody the crux of the problem. Instead, the issue lies with a tendency to examine singular variables. According to a standard reductionist approach, past experimental studies have isolated just one factor by eliminating all “extraneous” variables. In one study of the English article, for example, two types of CF were examined, along with language aptitude, by separating treatments into three groups (DCF only group, DCF and MCF group, and a control group). Although results correctly isolate treatment variables, showing that DCF was most effective.

1SUNY Korea, Incheon, South Korea

Corresponding Author:
Andrew Schenck, SUNY Korea, 119 Songdo Moonhwa-Ro, Incheon 21985, South Korea.
Email: Schenck@hotmail.com
effective when provided with MCF (Sheen, 2007), the effect is very context specific. Since differences in variables like target feature (Van De Guchte et al., 2015; Varnosfadrani & Basturkmen, 2009) and learner characteristics (Ellis, 2010) are known to influence corrective feedback, the same results can only be reliably obtained when the experimental environment is replicated. Participants from this study came from various different ethnic and linguistic backgrounds. Not only is it difficult for both researchers and teachers to recreate this classroom environment to repeat the effects, combined evaluation of learners from different backgrounds sends an implicit message that “one-size-fits-all” solutions are needed. Not surprisingly, research conducted with a new target feature (the hypothetical conditional) and more linguistically uniform learners, all of whom had the Japanese L1, yielded contradictory results (Shintani et al., 2014).

Past experimental studies of CF are expertly designed, yet they attempt to isolate variables in a scientific tradition that does not accurately characterize the real world. These experimental studies cannot be adapted to the classroom, primarily because factors that impact learning are codependent. Due to a need for more holistic understanding of multiple variables, researchers have called for additional qualitative research (Lee, 2019b). Ultimately, such calls reflect a greater need for comprehensive study of learner differences, an area of research that has been underexplored (Li & Vuono, 2019). If a more holistic understanding of CF is to be obtained, discrepancies in past research must be clarified, allowing for better prediction of learner outcomes. Like qualitative research, meta-analysis may be useful in providing a more comprehensive perspective. This technique, which collates data from several experimental studies, provides new insights about how multiple variables are related. By using meta-analytic investigation to gain a new perspective of CF, educators may finally obtain the knowledge needed to adapt theory to practice.

**Literature Review**

Before the effect of any pedagogical technique can be fully understood, we must first comprehend the causes. Analysis of past research suggests that type of grammatical feature is one “possible contributor” to the impact of CF (Li & Vuono, 2019, p. 103). In some ways, this “possible” contribution seems to be an understatement. It equates to the idea that “what you teach is a possible contributor to what students learn.” Although choices of grammatical feature would seem to have a significant impact, this idea is not clearly established through past research, which often yields variable results. Failure to concretely understand the role of grammar has compelled researchers to turn toward background variables, which focus on the mediating effects of student and teacher beliefs (Bao, 2019; Couper, 2019; Han, 2017). As in the case of earlier studies, this research does not provide the holistic perspective needed to consistently use different types of CF.

Earlier research does acknowledge that grammar type influences the effectiveness of CF. Unfortunately, this research has used overly simplistic categorization to analyze grammatical differences, thereby limiting understanding. Ferris (2006) first separated grammar into just two groups: systematic features (e.g., past regular tense or English article), which have rules that can be taught, and lexical features (e.g., the past irregular tense or collocations), which lack “teachable” rules. Although the untreated distinction of lexical features has now been debunked (Van Beuningen et al., 2012; Wang & Jiang, 2015), oversimplification of grammatical features continues to be a problem within new studies, which consistently use a simple binary dichotomy such as simple versus complex or early versus late (Spada & Tomita, 2010; Van De Guchte et al., 2015; Varnosfadrani & Basturkmen, 2009). Modern research generally recognizes the importance of grammatical differences, yet experimental designs continue to reflect the traditional bias. A recent study of CF with metacognition, for example, examined only two features: the third person singular (-s) and possessive determiner (his/her) (Sato & Loewen, 2018). Both of these features reflect past distinctions, arbitrarily dividing grammar into systematic and lexical categories. As language teachers often realize, grammar is not so simplistic. It is semantically, morphosyntactically, and phonologically variable. By cultivating a “one-size-fits-all” paradigm for grammar, the diversity of these features and their influences on corrective feedback remains enigmatic.

**Factors of Grammatical Difficulty: Morphosyntactic, Semantic, and Phonological Complexity**

In reality, characteristics of grammatical features differ significantly, affecting how they are acquired. Features that are often characterized as “early,” such as the irregular past, plural -s, and definite article (Varnosfadrani & Basturkmen, 2009), vary in phonological salience, regularity of form, and semantic complexity. The past irregular tense, for example, has sonorant vowel sounds, making it easier to hear, whereas the plural -s is less phonologically noticeable, often lacking a vowel (e.g., parents). Concerning regularity, the plural -s is highly systematic in form, while the irregular past has many different lexical variants, which require a more sophisticated understanding of form-meaning mappings. Concerning the definite article, it is phonologically salient and regular in form, yet has a heightened semantic complexity than the other “early” features. It may be used to signify general cultural use (e.g., the sun), immediate situational use (e.g., Don’t go in there. The dog will bite you!), perceptual situational use (e.g., Pass me the salt.), and local use (e.g., the car/the pub) (Celce-Murcia et al., 1999).
Grammatical features that are often characterized as “late,” such as the indefinite articles a and an, regular past, relative clause, passive voice, and third person -s (Varnosfadrani & Basturkmen, 2009), are also more diverse than a simple binary dichotomy would suggest. Like the plural -s feature, third person -s is systematic and morphologically simple, yet challenging to hear or perceive. Unlike the plural -s, third person singular morphology links a subject with the predicate verb (e.g., He eats). This cognitive link between agent and action appears to explain why this feature is so difficult to acquire (Pienemann, 2005). In the case of the passive voice, constructions like, He was killed, include not only regular past tense features in the verb (e.g., killed), but a lexically contrived, past irregular component (e.g., was). Relative clauses require a combination of words in a syntactic order, as well as correct conjugation of its constituents. Such characteristics make the feature easy to perceive, yet morphosyntactically and semantically complex. Overall, review of grammatical features reveals that simple binary dichotomies are insufficient as a means to explain variability.

Although grammar differs in a number of ways, impact of this variability on the effectiveness of CF has never been fully investigated. Because MCF provides explicit information about grammar or meaning needed for correct use of a target feature, it may be most effective when used with grammatical features like articles, which are semantically complex. Oral forms of CF like recasts, which provide a “reformulation of all or part of a student’s utterance, minus the error” (Lyster & Ranta, 1997, p. 46), may be more effective than MCF or DCF when used with phonologically complex grammatical features. As the past regular tense has only one orthographic form and meaning (ed is used to mean “past”), yet three phonological variants (“t,” “d,” or “id”), it may benefit more from recasts, which provide valuable phonological input. There may be times when a specific kind of feedback is needed for a target feature. Unfortunately, little research has been conducted to examine how different types of CF are affected by unique semantic, morphosyntactic, and phonological characteristics of each grammatical feature. This failure may have limited our understanding of how and when to use CF.

**Additional Factors of Grammatical Difficulty:**

**Learner Proficiency and L1**

Although morphosyntactic, semantic, and phonological complexity are all important, learner characteristics like proficiency and L1 impact the acquisition of a grammatical feature. Both of these factors help to determine whether or not a target feature will be more difficult to acquire. Studies of the Processability Theory suggest that grammatical features are more easily acquired at distinct stages of proficiency, which move from morphosyntactic operations within a phrase (a noun, verb, or adjective phrase) toward more sophisticated operations that link different phrases or clauses (Pienemann, 2005). English learners first acquire SVO sentences and basic aspects of morphology like the plural -s. They then acquire aspects of do support, question inversion, and phrasal verb use, all of which reflect cognitive links between multiple noun and verb phrases. Finally, learners begin to use features such as the hypothetical conditional and embedded questions (e.g., Could I ask what the trouble is?), which require complex dependent clauses (Mohamed Salleh et al., 2020; Pienemann, 2005). Because learners become able to acquire grammatical features at a specific level of proficiency, timely introduction of CF may be a crucial determinant of efficacy. This notion is supported by past research of explicit grammar instruction, which is more effective when the target feature is just above a learner’s ability level (Dyson, 2018; Dyson & Håkansson, 2017; Pienemann, 1989).

In addition to proficiency, a learner’s L1 influences how difficult a grammatical feature is to acquire. This influence is revealed by a recent study, which shows that explicit focus on L1 processing routines can be used to improve accuracy and fluency in the L2 (McManus & Marsden, 2019). Research suggests that differences in L1 morphology (free vs. bound morphemes) or syntax (head-final vs. head-initial languages) have a negative impact on the process of acquisition (Luk & Shirai, 2009; Maleki, 2006; Shin, 2015). Conversely, recent meta-analysis of Korean and Chinese learners suggests that similarities of L1 linguistic structures make an L2 target feature easier to learn (Schenck, 2020; Yang, Cocq et al., 2017). While such research is insightful, a primary focus on Asian languages has limited understanding of L1 influence. Because Indo-European languages like Persian include grammatical features that are not present in languages like Korean or Chinese (e.g., an article system), further investigation may allow for comparison, leading to a more holistic perspective on the efficacy of CF.

Because grammatical difficulty cannot be isolated into a single variable, multiple factors must be simultaneously studied. The challenge posed by a specific grammatical feature can only be ascertained through examining the complexity of a grammatical feature (morphosyntactic, semantic, and phonological) and learner characteristics (learner proficiency and L1). Through comprehensive analysis of both grammatical and learner variables, difficulty posed by specific target features may finally be understood, allowing educators to predict the impact of CF techniques. Further study of grammatical difficulty may allow educators to better understand how CF will impact the learner, leading to better choices that promote acquisition.

**Types of Corrective Feedback**

A variety of CF types such as recasts, clarification requests, elicitation, metalinguistic feedback, indirect feedback, and direct feedback may be used to increase grammatical accuracy and foster acquisition (Lyster & Ranta, 1997; Tedick &
De Gortari, 1998). While these techniques all compel a learner to correct an error, they differ according to how much information is provided. Both elicitation and clarification requests ask the learner to fix an error without providing any explicit information about the mistake (e.g., Excuse me, could you repeat that?). Metalinguistic and indirect feedback also prompt the learner to correct an error, yet they provide explicit information for support. Concerning indirect feedback, errors are underlined or highlighted. Metalinguistic feedback provides explicit feedback in two main forms, through coding errors (within writing) or providing grammatical hints without a correction (e.g., Should you use the past tense or present tense in that sentence?) (Sanavi & Nemati, 2014). While clarification requests and elicitation differ from metalinguistic and indirect feedback in the degree to which explicit information is provided, all of the techniques compel a learner to fix a grammatical error without providing the correct form (see Table 1). It is the learner that must reformulate their speech or writing. In contrast to these types of CF, which are collectively called prompts, techniques such as recasts and direct feedback rely on the instructor to reformulate an error. Recasts, which generally provide a correction orally, do not explicitly suggest that a learner has made an error, whereas direct feedback, which is usually delivered in writing, provides information that can be explicitly processed and reviewed. Although recasts are implicit and direct feedback is explicit, both techniques provide a revised and correct form to students. Thus, they are termed reformulations (see Table 1).

There are four different characteristics of corrective feedback: prompt, recast, implicit, and explicit (see Table 1). Each one of these characteristics may influence efficacy of a CF technique. In the case of explicit CF, increased conscious awareness can help the learner, providing information that promotes error correction. At the same time, increased conscious awareness also increases cognitive load, which may decrease accuracy. If a learner is attempting to use a difficult grammatical feature (as determined by proficiency level and L1), this learner’s ability to process will decline. Without providing CF that is manageable, learners may become “unable to process feedback effectively and may experience something akin to the learning breakdown predicted by cognitive load theory” (Hartshorn et al., 2010, p. 88). In the case of implicit prompts, no real assistance to correct an error is given, leaving a learner with the highest burden. Although different forms of CF may vary in the cognitive load they place on the learner, effects of these differences have not yet been studied. Concerning written corrective feedback, for example, the links between working memory (cognitive load) and CF type are currently unknown (Li & Roshan, 2019).

In addition to traits outlined in Table 1, medium of CF delivery, either written or oral, must also be considered. As pointed out by Li and Vuono (2019), all written forms of feedback are explicit since a learner is able to view grammatical emphasis over and over. Oral prompts and reformulations, however, do not provide the same explicit scaffold. They force a learner to process an error in real time. In some way, oral input provided by recasts can serve as a scaffold for phonologically challenging grammatical features. This view is supported through research of structured input, which has been shown to be useful for the past regular -ed, morphology that is phonologically complex, yet semantically and syntactically simple (Benati & Angelovska, 2015; Benati & Batziou, 2019).

Grammatical complexity, proficiency, L1, and CF type all exert an influence on the process of grammatical acquisition. Past studies have tried to isolate these variables one by one, carefully designing experiments that eliminate the influence of “extraneous” factors. Because each factor is actually codependent, discreet isolation is impossible. To better clarify relationships between multiple influences, more comprehensive study is needed. Such study will finally provide the holistic perspective needed to adapt theory to practice. Essentially, understanding interconnected influences of grammar acquisition will make outcomes of CF predictable, thereby giving educators the knowledge needed to make important choices about how and when to use such techniques.

### Research Questions

To provide a more holistic understanding of the multiple influences that impact CF, the following questions were posed:

1. Which kind of CF is most effective for promoting accuracy in the production of speech or writing?
2. Does the efficacy of CF differ based upon the target feature? What types of CF are most effective with each grammatical feature?
studies needed to include the following:

- Instruments were excluded.
- Studies which used multiple choice or fill-in-the-blank were also included in the meta-analysis. Being an oral measure, pressure is placed on the learner for an immediate response, which prevents conscious correction.
- All of the tests required either spoken or written production.

Method

The present study was designed to examine the impact of both grammatical (semantic, morphosyntactic, and phonological) and learner (proficiency and L1) characteristics on the efficacy of CF. In order to obtain data for analysis, a search was conducted for experimental studies which included information about the target feature, learner proficiency, and L1. To avoid confusion over linguistic variability of participants, studies had to include learners with only one L1, Persian (Farsi). After studies were located and statistically analyzed, results were then compared to a previous meta-analysis of CF which included only Korean learners (Schenck, 2020).

Google scholar was used to locate all research for initial review. The database includes not only journals from Scopus and Web of Science, but sources outside these two databases, making it a comprehensive resource for existing research. The database was systematically searched by first using general search terms *Iranian* and *feedback*. Following this, these search terms were used with specific grammatical features *(plural, past tense, past regular, past irregular, passive, third person, questions, article, definite article, indefinite article, phrasal verb, verb particle, conditional)* and feedback types *(feedback, recasts, indirect feedback, direct feedback, metalinguistic feedback, prompts)*.

Explicit (conscious) knowledge of a grammatical feature was not examined in this study, primarily because it does not always equate to increased accuracy in natural speech or writing. To help ensure that implicit (unconscious) knowledge was being examined, studies had to include pretest and posttest measures with the following qualities for the measurement of implicit knowledge: the communication of ideas, not rules; a focus on meaning not form; and the avoidance of using metalanguage (Ellis, 2009). Studies that used oral imitation were also included in the meta-analysis. Being an oral measure, pressure is placed on the learner for an immediate response, which prevents conscious correction. All of the tests required either spoken or written production. Studies which used multiple choice or fill-in-the-blank instruments were excluded.

In order to be considered for the present meta-analysis, studies needed to include the following:

1. A type of CF (including time for treatment and methods of delivery)
2. Pretest and posttest measures of production (either oral or written)
3. Information about the type of grammatical feature targeted
4. Information about proficiency
5. Participants that had a Persian (Farsi) L1

In addition to these criteria, all of the studies were reviewed for quality, ensuring the selection of studies with sound experimental methods which included adequate description for replication. All of the studies were composed in English and most had from one to four treatments, with the exception of Davaribina and Hossein Karimi (2014), Ghoorchaei and Gharanjik (2020), and Zarei et al. (2018), which received from 6 to 15 sessions of treatment (see Appendix A for more information). The length of treatment session was qualitatively considered as effect size was analyzed. Potential influences of treatment length were considered in the limitations section. Of the 374 files discovered, only 11 met all the criteria necessary for analysis (see Appendix A for more information on these studies).

Corrective Feedback

Types of CF were analyzed according to four categories suggested in a study by Lyster and Saito (2010): implicit prompts (clarification requests and elicitation), implicit reformulations (recasts), explicit prompts (metalinguistic and indirect feedback), and explicit reformulations (direct feedback). Specific types of feedback were charted based upon whether or not they were written or orally delivered. As suggested by Li and Vuono (2019), all written forms of feedback carry an explicit quality, making this separation necessary.

Grammatical Features

After studies were collected and examined, there was a total of 37 treatment groups from 11 studies (see Appendix A). The following grammatical features were available for analysis: indefinite and definite articles (four studies with fourteen groups), regular past (one study, four groups), irregular past (one study, four groups), passive voice (one study, four groups), wh questions (one study, two groups), future tenses (one study, two groups), and conditional (one study, one group). A study by Zarei et al. (2018) combined past (past simple and past progressive) and future tenses (“be going to” and present progressive for future) together for treatment of four groups, while the study by Davaribina and Hossein Karimi (2014) combined articles with the past simple tense for treatment of two groups. These studies were included to show how heightened complexity of multiple features impacts CF.

Proficiency Level

Studies selected for meta-analysis included the following proficiency levels: elementary, pre-intermediate, low-intermediate, and upper intermediate. Most of the studies were at the intermediate level (see Appendix A). A study by Rassaei (2013) included only information about prior English
A study by Khezrlou (2019) assessed learners at the CEFR B2 level, which may be considered to be upper intermediate. Because proficiency levels were determined via several different measures, results of any comparison between studies must be interpreted with caution. Liu and Brown (2015) have cited that assessments of proficiency used in past studies are often unsystematic or neglected. Despite such methodological issues, consideration of proficiency level from past studies may help in identifying some key relationships with other variables (CF type, grammar type, and L1), which will heighten our understanding of how and when to use CF.

**L1 Similarity**

Effect sizes were analyzed based on an English target feature’s similarity or dissimilarity to the Persian L1. Persian and English have many grammatical similarities, yet notable differences also exist, which may complicate the effectiveness of CF instruction. The two languages were systematically compared by examining the following traits: presence/absence of a target feature; free, bound, or lexical morphological attributes; and word order (head initial vs. head final). Differences in these characteristics have been shown to influence the process of acquisition (Luk & Shirai, 2009; Maleki, 2006; Shin, 2015).

Unlike English, which is an SVO language, Persian is an SOV language. Questions use the same SOV word order (subject, interrogative pronoun, and verb) as in the following example:

انتست چی این
Is what this

The above question, which is read from right to left (“This what is?”), uses declarative sentence structure, without any inversion. There is also no do support in Persian. Instead of changing word order, intonation is used to suggest a question (North Carolina State University, 2013, p. 5).

Concerning verb tense, Persian uses a past tense verb that is completely different from the present form (similar to lexical past in English). To correctly use the past tense, a morphological ending that reflects the subject (I, You, they, etc.) must also be added, with the exception of third person, which does not require an ending (North Carolina State University, 2013, p. 9). Like the lexical past tense, “the present stem has very little or no resemblance to the infinitive form of the verb and needs to therefore be learned through repeated exposure to the print” (North Carolina State University, p. 19). Concerning the future tense, it “is used almost exactly like the English future tense; the only difference being that it is also very common in Persian to use the present tense for expressing future actions” (Mazdeh, 2013, para. 1). The future tense is constructed by using the present simple tense form of “to want” (خواستن) with a past tense stem of the main verb. This structure somewhat resembles English structures that use an auxiliary “going to” with a consistent form of main verb (“to eat”), while verb forms are lexically different (Mazdeh, 2013).

Concerning conditionals, the if structure is mainly initial in imperative and declarative statements (Abdollahi-Guilani et al., 2012). The Persian word for if (اگر) is generally used at the beginning of the conditional clause, followed by a main clause which uses the future tense. Despite Persian being a head-final language, the if marker appears at the beginning of the conditional clause, as in English. Grammar structure is also very similar to English.

Both English and Persian have an article system. Unlike English, however, nouns do not carry a definite morpheme as in the English “the.” Instead, definiteness is implied through the lack of an indefinite article. Within colloquial speech, a bound morpheme (-e) may be attached at the end of a noun (in accordance with the head-final nature of the language) to signal definiteness (Momenzade & Youhanaee, 2014).

The final feature analyzed within this study, the passive, has some notable dissimilarities with English. The feature also exists in Persian, yet the auxiliary verb appears at the end of the sentence as in “apples (سیب) are washed (شسته)” (Li & Roshan, 2019, p. 5).

**Effect Size**

To calculate effect size, pretest scores (M2), posttest scores (M1), and standard deviations (SD2 and SD1) were inserted into Cohen’s d formula for effect size (Spada & Tomita, 2010, p. 307):

\[
d = \frac{M1 - M2}{\text{SQRT} \left[ \frac{SD1SD1 + SD2SD2}{2} \right]}
\]

Initially, calculating Cohen’s d by comparing posttests of treatment groups with those of control groups was considered. Four of the studies, however, did not use a control group (they studied different types of treatment). Using the pretest/posttest calculation allows for inclusion of more grammatical features and provides an accurate measure of improvement for each treatment group. While pretest/posttest scores were used for this study, calculation of Cohen’s d using posttests from treatment and control groups was performed to the check results. Of the studies with control groups, the Cohen’s d using posttests from treatment and control groups resembled the Cohen’s d using pretest and posttest scores. Effect sizes for grammatical features and CF types differed in nearly the same proportions for both calculations.

After effect sizes were calculated, they were graphically charted according to grammatical feature and type of CF. Since the meta-analysis was limited to only 11 studies, there was not enough data to collate effect sizes based on L1 differences or proficiency. Instead, these factors were qualitatively considered for each grammatical feature by looking at characteristics of individual research studies responsible for
the effect size. Potential differences like treatment length were also considered. Such qualitative inquiry helped to ensure that complex interaction between factors was identified, thereby providing a more holistic perspective of CF. Such analysis also helped to address issues of heterogeneity between studies by acknowledging and interpreting different influences from individual studies. Following analysis, potential problems with meta-analytic comparison of studies was also considered, along with a need for methodological changes within future research.

Results

Contrary to prior meta-analysis suggesting higher gains for explicit emphasis of grammar (Spada & Tomita, 2010), implicit prompts (clarification requests) had the highest overall effect size (d = 3.40). Explicit forms of written DCF (d = 2.98), oral MFC (d = 2.26), written indirect feedback (d = 1.80), and written MFC (d = 1.47) had the next highest effect sizes. Finally, implicit recasts (d = 1.50) had the lowest effect sizes (see Appendix B for more information). Overall, written forms of CF (d = 2.21) had a slightly larger effect size than oral forms of CF (d = 2.09). This finding may be explained by the characteristics of written CF, which is explicit and more highly scaffolded (learners have time to process and correct an error). While all of the collated results are interesting, separation of analysis by target feature and learner characteristics (proficiency and L1) suggests that general claims about the efficacy of CF are misleading (see Figure 1). Context-specific variables concerning grammar and learner background must be considered before a CF technique can be labelled either effective or ineffective.

Effect sizes for CF varied considerably based on the target feature (see Appendix C). This variance appears to be explained by grammatical difficulty, as defined by characteristics of the target feature (morphosyntactic, semantic, and phonological) and learners (proficiency and L1). To illustrate this point, the overall effect size for clarification requests may be examined. This statistic is based upon just one study by Khezrlou (2019). On the surface, it suggests that Persian learners benefit more from implicit prompts, which parallels findings in an experimental study of Chinese learners (Yang & Lyster, 2010). However, the broad claim that implicit prompts are more effective cannot be asserted, primarily because the Persian and Chinese learners in both studies received CF for the same grammatical feature (past simple tense). Additionally, participants within both studies appear to have had a larger degree of English proficiency than the average learner. Whereas the Chinese participants were sophomores majoring in English language and literature, Persian participants were learners from a private institute who scored at the B2 level of the Common European Framework of Reference (CEFR). Using implicit prompts at a higher proficiency level for the past simple tense may explain the results. For Korean learners, results were reversed, with recasts revealing a larger effect size (Schenck, 2020). Learners in these studies, however, were deemed to be pre-intermediate or low-level (Cho, 2012; Kim & Cho, 2017). Perhaps the cognitive burden placed on low level learners by implicit prompts explains the lower performance for Korean learners.

For studies which collectively examined past or future tenses, recasts had the highest effect sizes (see Figure 1). Like prior research conducted in Korean contexts, Persian studies of past verb tenses used low intermediate and pre-intermediate learners (Hamtaei et al., 2015; Zarei et al., 2018). Replication of results among low-level learners, who are more heavily influenced by recasts, suggests that proficiency influences what type of CF is best for a target feature. Being at lower proficiency levels, learners may have more difficulty self-repairing, explaining why a scaffolded reformulation is needed. Oral MFC had a sizeable effect size, yet less so than recasts. Oral MFC was also least effective for the past tense, which has a systematic morphological element (-ed) that does not exist in Persian. L1/L2 dissimilarity may have limited the efficacy of oral MFC, while simultaneously promoting the impact of recasts, which provide valuable input concerning three phonological variants. Perhaps oral MCF would be more effective at a slightly higher proficiency level when challenges posed by tenses are less substantial. Overall, results of CF used with verb tenses suggest that both proficiency and L1 similarity (or dissimilarity) define grammatical difficulty, thereby influencing the efficacy of different types of feedback.

Regarding the conditional feature, effect size for metalinguistic feedback was lower than that used for other grammatical features (d = 1.27). Students in the Persian study examined (Ghoorchaei & Gharanjik, 2020) were deemed to be beginner learners, which may explain why this more morphosyntactically and semantically complex feature had a lower effect size. For Korean learners, effect size of treatment for the conditional feature was over double (d = 2.67). Participants in this study, however, were at a higher, intermediate proficiency level, which may explain the findings (Kim, 2019). Using metalinguistic feedback for the conditional feature may be more appropriate at an intermediate level. The head-initial characteristic of the conditional clause in Persian is also similar to that in English, whereas Korean has a head-final marker. Metalinguistic emphasis of this L1/L2 difference may be more necessary in Korean, yet not in Persian, explaining the findings.

Effect size for the article (a and the) reveals that metalinguistic feedback is superior to recasts for Persian learners, who were at intermediate levels of proficiency (see Appendix A). This finding agrees with other studies conducted in Korea, which suggest that metalinguistic feedback may provide information that assists the learner to understand more complex semantic relationships associated with this feature (Schenck, 2020). At the same time, the difference between
feedback types is small (difference of 0.17) compared to that of Korean learners (difference of 0.60), which suggests that recasts are more effective with Persian learners. Whereas the Persian language has an article system, Korean does not. A relatively large effect for recasts may reflect the presence of an article system in the L1, which promotes acquisition for the Persian learner through a transfer effect. For Korean learners, the article system does not exist, suggesting that implicit techniques may not be enough to help learners recognize an error.

Metalinguistic learning was highly effective for wh questions, revealing a much larger effect size than recasts for the intermediate learners that were studied (Figure 1). The larger difference in effect between these two feedback types (difference of 1.28) may reflect both grammar type and proficiency level. According to the Processability Theory, wh questions are acquired at intermediate stages, before more complex clauses (e.g., conditionals). Learners in the study of wh questions (Rassaei, 2014), being at an intermediate level, may have been cognitively ready to benefit from explicit emphasis of the target feature. The use of an explicit prompt (MFC) may help learners at the right proficiency to recognize differences between the L2 and Persian.

When both articles and simple past were emphasized within one study (Davaribina & Hossein Karimi, 2014), more highly scaffolded (yet explicit) DCF was more useful than indirect corrective feedback. In another study (Li & Roshan, 2019, p. 5), DCF was also more useful than MFC. DCF had an overall higher impact than both MFC and indirect corrective feedback, both of which are explicit prompts. The scope and complexity of target features may explain why scaffolded reformulations, in the form of DCF, may have been more effective. Emphasizing both articles and the simple past increases the scope, thereby making the target of emphasis more complex. The passive structure is also highly complex, including not only L1/L2 differences in word order, but morphological and lexical elements that increase scope and complexity. DCF provides a written reformulation that lessens pressure, allowing learners time to examine challenging target features.

Results suggest that type of grammatical feature, as well as associated learner variables such as L1 similarity or proficiency, influence the efficacy of different CF types. Low levels of proficiency, L1/L2 disparities, or heightened scope of grammatical emphasis all seem to increase the difficulty of acquiring a grammatical feature. As complexity increases, learners appear to benefit from CF that serves as a scaffold (both implicit and explicit reformulations). Recasts and DCF are more effective for intermediate learners when used with target features that are larger in scope (e.g., combined verb tenses/combined article and past tense) or more complex (e.g., passive tense).

Figure 1. Mean effect sizes based on type of grammatical feature and CF.
In contrast, when a grammatical feature is easier for the learner (as determined by proficiency and L1 similarity), prompts appear to be more effective. MFC yielded higher effect sizes when used with both English articles and wh questions. These features are easier for intermediate learners to acquire than the conditional structure, which requires creation and manipulation of larger clauses (according to the Processability Theory). With this in mind, it is not surprising that MFC for the conditional feature yields lower effect sizes with elementary learners. Implicit prompts (clarification requests) were most effective when targeting the past simple tenses. Past regular and irregular tenses are semantically, morphosyntactically, and phonologically less complex than other features like the English article, wh questions, or conditionals. The past regular tense is more systematic in form (only -ed) than the more semantically complex article with three variants (and, a, or the). The past irregular tense is similar to Persian (uses a lexical variant of the present tense for each verb), explaining why implicit prompts were even more effective for this feature. Overall, grammatical features that are easier for a learner to acquire appear to benefit more from prompts.

Results of this study are intriguing. They suggest that grammatical difficulty determines the effectiveness of CF type. Although further research is needed, a relationship between grammar and CF appears to have emerged (see Table 2).

As complexity increases (determined by grammatical characteristics and learner background), more scaffolded forms of CF (explicit feedback or reformulation) may be needed. Conversely, when the scope or complexity of a grammatical feature decreases, less scaffolded forms of CF (implicit feedback or prompts) may be required. Although assertions about CF are supported by evidence obtained from the present meta-analysis, along with cross-linguistic comparison of both Chinese and Korean learners, more research is needed.

### Conclusion

Results of the present study reveal that different types of CF cannot be understood without grasping the difficulty of a target feature, which is defined by qualities of the feature itself, along with characteristics of the learner (proficiency level and L1). Because traditional experimental methods isolate these variables, relationships are effectively hidden, limiting our understanding of how to use CF in realistic contexts.

Through meta-analysis, relationships between multiple influences of CF become clearer. As grammatical difficulty increases, recasts and direct feedback appear more useful as a kind of scaffold. Learners are freed from the burden of reformulating an error on their own. In cases where grammatical features of the L2 differ from those in the L1, explicit techniques appear to be effective. Explicit focus on differences appears to have a large impact. Finally, learners at a proficiency level that is high enough to focus on a specific target feature (meaning they have adequate cognitive resources) may benefit from prompts, which place a burden on the learner to reformulate an utterance.

### Limitations and the Need for Future Research

While information obtained from this meta-analysis was insightful, experimental research which can be used for comprehensive analysis of CF is limited. Very few studies include all of the variables needed to assess grammatical difficulty (characteristics and scope of a grammatical feature; L1 similarity or dissimilarity; and learner proficiency). In addition to limitations associated with assessment of grammatical difficulty, studies do not often examine a range of different CF types and grammatical features. Concerning implicit prompts (clarification requests), for example, only one study was available for meta-analysis of Persian learners. This study, like others used to examine both Chinese and Korean learners, chose the same grammatical feature (the simple past tense). The tendency to use the same grammatical features with the same types of CF has limited our understanding. In reality, all four types of CF (implicit prompts, explicit prompts, implicit reformulations, and explicit reformulations) need to be tested with a range of grammatical features, and with learners of differing L1s and proficiency levels. More comprehensive study will give us concrete information concerning how to adjust CF techniques as grammatical difficulty either increases or decreases.

Variable treatments and assessment of implicit knowledge are also problematic in past studies of CF. Regarding treatments, for example, studies that combined different grammatical features (Davaribina & Hossein Karimi, 2014; Zarei et al., 2018) also had the highest number of treatments, from 6 to 10. In this case, realizing the role of treatment length may be problematic, since other variables like grammatical

### Table 2. Relationship Between CF and Grammatical Difficulty.

| Type of CF           | Degree of Assistance | Difficulty of Grammatical Feature |
|----------------------|----------------------|----------------------------------|
| Implicit prompt      | Less scaffolded      | Lowest                           |
| Explicit prompt      |                      |                                  |
| Implicit reformulation |                    |                                  |
| Explicit reformulation | More scaffolded     | Highest                          |

...
difficulty, CF type, or learner characteristics may all have influenced the results. In the case of the conditional feature, however, which had an extensive daily treatment over 3 weeks (Ghoorchaei & Gharanjik, 2020), lower effect size would suggest that grammatical feature type and proficiency level played a larger role than treatment length. No matter how many small-scale individual experimental studies are conducted, using different methods (e.g., treatment lengths, tasks, or assessments) threatens the validity of any comparison. While all effort was made to eliminate these threats to validity in the present meta-analysis, results require further experimental verification. More large-scale studies that systematically utilize and assess the impact of CF are clearly needed.

To better understand multiple variables involved in CF, Lee’s (2019b) argument for more qualitative research seems to be a valid one. It would allow for consideration of multiple variables in relation to learner background. Unfortunately, this type of research suffers from the same weaknesses of past experimental studies. Results obtained from different learners cannot be collated or compared if treatments and measures are not systematically utilized. It is important to conduct studies that carefully record relevant variables for CF delivery in systematic ways. For example, an individual’s proficiency level, L1, and cognitive capacity (working memory) can be carefully assessed according to the same instruments. Conducting several studies in this way will create a corpus of experimental studies, which may be systematically searched.

While problems with past experimental studies is clear (Liu & Brown, 2015), the best way to develop future research is more uncertain. Essentially, teachers need a guide to use CF techniques, which considers both grammar type and learner background. With this in mind, more research is needed to identify when each form of CF should be introduced. Through understanding the time in which more scaffolded techniques are needed (or not needed), teachers may be able to tailor instruction to learner needs, thereby maximizing the effectiveness. Far from being finished with study of CF, we have only scratched the surface. Either large-scale studies of multiple variables or small, collatable mini studies must be conducted to provide a holistic perspective that teachers can really use.

### Appendix A. Studies Included in Meta-Analysis.

| Study | Proficiency level | Treatment | Grammatical feature | Pretest/posttest type | Average Cohen’s $d$ |
|-------|-------------------|-----------|---------------------|----------------------|---------------------|
| Rassaei (2020) | Intermediate | Four treatments in 8 days | Definite the, and indefinite a | Picture description task (oral) / story writing task | 1.8887 |
| Khezrlou (2019) | CEFR B2 | One treatment session on day 8 | Regular and irregular past | Oral narratives | 1.8152 |
| Rassaei (2013) | 26 months English instruction | Two treatments in 1 week | Definite the, and indefinite a | Written narrative with picture sequences | 2.8086 |
| Rassaei et al. (2012) | Lower intermediate | Two treatment sessions each | Definite the, and indefinite a | Elicited oral imitation test (EOIT) | 1.0803 |
| Rassaei (2015) | Upper Intermediate | Three treatments in 1 week | Definite the, and indefinite a | Oral and written narrative with pictures | 1.7824 |
| Rassaei (2014) | Intermediate | Three sessions over 3 days in a week | Wh questions fronting, auxiliary verb inversion, and do-support | Oral production test / spot the differences in two pictures | 1.7602 |
| Hamtaei et al. (2015) | Low Intermediate | One session 30 minutes or less - 1 day | Future tense “will” or “going to” | Written paragraph about plans for future | 2.0977 |
| Davaribina and Hossein Karimi (2014) | Intermediate | Six sessions total | Articles and simple past | Written narrative with pictures | 3.6098 |
| Zarei et al. (2018) | Pre-intermediate | 10 sessions over 5 weeks about 75 minutes | Simple past, past continuous, and future | Oral description of picture strips | 3.6530 |
| Ghoorchaei and Gharanjik (2020) | Elementary | 3 weeks is all the info given | Hypothetical conditional | Writing task on “general topic of interest” | 1.2725 |
| Li and Roshan (2019) | Intermediate | 10 minutes treatment in week 2/one session | English passive voice (are washed apples) | Recreated text based on process diagram | 1.7428 |
Appendix B. Overall Effect Size Based on Feedback Type.

| Feedback Type                  | Mean    | N   | Std. deviation |
|-------------------------------|---------|-----|----------------|
| Direct feedback (written)     | 2.9757  | 3   | 1.68067        |
| Indirect feedback (written)   | 1.7976  | 2   | 0.74771        |
| Metalinguistic feedback (written) | 1.4688  | 2   | 0.00235        |
| Metalinguistic feedback (oral) | 2.2647  | 10  | 1.20408        |
| Recast (oral)                 | 1.4968  | 15  | 1.12530        |
| Prompt (oral)                 | 3.3963  | 4   | 0.45330        |
| Total                         | 2.1116  | 37  | 1.25693        |

Appendix C. Overall Effect Size of CF Based on Each Grammatical Feature.

| Target Feature               | Mean    | N   | Std. deviation |
|-------------------------------|---------|-----|----------------|
| Regular past                 | 1.6200  | 4   | 1.60751        |
| Irregular past               | 2.0104  | 4   | 2.04723        |
| Conditional                  | 1.2725  | 1   | 0.00000        |
| Article                      | 1.8439  | 14  | 0.97946        |
| Past simple                  | 3.3281  | 2   | 0.44218        |
| Wh question                  | 1.7602  | 2   | 0.90472        |
| Future tenses                | 3.0378  | 4   | 1.28436        |
| Articles and simple past     | 3.6098  | 2   | 1.81518        |
| Passive                      | 1.7428  | 4   | 0.38015        |
| Total                        | 2.1116  | 37  | 1.25693        |

Declaration of Conflicting Interests
The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author received no financial support for the research, authorship, and/or publication of this article.

Ethics Statement
This study did not include human subjects or animals. Therefore, this section is not applicable.

ORCID iD
Andrew Schenck https://orcid.org/0000-0002-3864-6267

References
Abdollahi-Guilani, M., Yasin, M. S. M., & Hu, T. K. (2012). A comparative analysis of conditional clauses in English and Persian: Text analysis. 3L: Language, Linguistics, Literature, 18(2), 83–93.
Bao, R. (2019). Oral corrective feedback in L2 Chinese classes: Teachers’ beliefs versus their practices. System, 82, 140–150. https://doi.org/10.1016/j.system.2019.04.004
Benati, A., & Angelovska, T. (2015). The effects of processing instruction on the acquisition of English simple past tense: Age and cognitive task demands. International Review of Applied Linguistics in Language Teaching, 53(2), 249–269. https://doi.org/10.1515/iral-2015-0012
Benati, A., & Batziou, M. (2019). The relative effects of isolated and combined structured input and structured output on the acquisition of the English causative forms. International Review of Applied Linguistics in Language Teaching, 57(3), 265–287. https://doi.org/10.1515/iral-2016-0038
Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. Journal of Second Language Writing, 14(3), 191–205. https://doi.org/10.1016/j.jslw.2005.08.001
Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1999). The grammar book: An ESL/EFL teacher’s course (2nd ed.). Heinle & Heinle.
Chen, S., Nassaji, H., & Liu, Q. (2016). EFL learners’ perceptions and preferences of written corrective feedback: A case study of university students from mainland China. Asian-Pacific Journal of Second and Foreign Language Education, 1(1), 1–17. https://doi.org/10.1186/s40862-016-0010-y
Cho, H. (2012). Relative effects of prompts and recasts on the development of implicit and explicit L2 knowledge. English Teaching, 67(4), 57–79. https://doi.org/10.1585/engtea.67.4.201212.57
Couper, G. (2019). Teachers’ cognitions of corrective feedback on pronunciation: Their beliefs, perceptions and practices. System, 84, 41–52. https://doi.org/10.1016/j.system.2019.04.003
Davaribina, M., & Hossein Karimi, S. (2014). The effect of different types of teacher written corrective feedback on Iranian EFL learners’ writing accuracy. Journal of English Language Pedagogy and Practice, 7(14), 23–39.
Dyson, B. (2018). Developmental sequences. In J. I. Liontas (Ed.), The TESOL encyclopedia of English language teaching (pp. 1–8). John Wiley & Sons.
Dyson, B. P., & Håkansson, G. (2017). Understanding second language processing: A focus on processability theory (Vol. 4). John Benjamins.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. International Journal of Applied Linguistics, 19(3), 221–246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Ellis, R. (2010). A framework for investigating oral and written corrective feedback. Studies in Second Language Acquisition, 32(2), 335–349.

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), Feedback in second language writing: Contexts and issues (pp. 81–104). Cambridge University Press.

Ghoorchaei, B., & Gharanjik, N. (2020). The impact of metalinguistic corrective feedback on Iranian EFL learners’ acquisition of the hypothetical conditional. AJELP: Asian Journal of English Language and Pedagogy, 8(2), 27–38. https://doi.org/10.37134/ajelp.vol8.2.3.2020

Hamtaei, B., Rahimi, R., & Branch, T. (2015). The effect of oral versus written error feedback on paragraph writing ability of Iranian low-intermediate EFL learners. Modern Journal of Language Teaching Methods (MJLTM), 5(1), 125–133. https://mail.mjltm.org/

Han, Y. (2017). Mediating and being mediated: Learner beliefs and learner engagement with written corrective feedback. System, 69, 133–142. https://doi.org/10.1016/j.system.2017.07.003

Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. TESOL Quarterly, 44(1), 84–109.

Khezrlou, S. (2019). Task repetition and corrective feedback: The role of feedback types and structure saliency. English Teaching & Learning, 43(2), 213–233. https://doi.org/10.1007/s42321-019-00025-2

Kim, J. H. (2019). Explicit written corrective feedback, language analytic ability, and L2 accuracy development. Korean Journal of Applied Linguistics, 35(2), 53–81. https://doi.org/10.17154/kjal.2019.6.35.2.53

Kim, S., & Cho, Y. W. (2017). The effects of recasts and working memory on Korean EFL learners’ past tense accuracy. English Teaching, 72(4), 105–132. https://doi.org/10.15858/engtea.72.4.201712.105

Lee, I. (2019a). Teacher written corrective feedback: Less is more. Language Teaching, 52(4), 524–536. https://doi.org/10.1017/s0261444819000247

Lee, I. (2019b). Teachers’ frequently asked questions about focused written corrective feedback. TESOL Journal, 10(3), 1–15. https://doi.org/10.1002/tesj.427

Li, S., & Roshan, S. (2019). The associations between working memory and the effects of four different types of written corrective feedback. Journal of Second Language Writing, 45, 1–15. https://doi.org/10.1016/j.jslw.2019.03.003

Li, S., & Vuono, A. (2019). Twenty-five years of research on oral and written corrective feedback in system. System, 84, 93–109. https://doi.org/10.1016/j.system.2019.05.006

Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. Journal of Second Language Writing, 30, 66–81. https://doi.org/10.1016/j.jslw.2015.08.011

Luk, Z. P. S., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles, and possessive a’s. Language Learning, 59(4), 721–754. https://doi.org/10.1111/j.1467-9922.2009.00524.x

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. Studies in Second Language Acquisition, 19(1), 37–66. https://doi.org/10.1017/s027263197001034

Lyster, R., & Saito, K. (2010). Oral feedback In classroom sla. Studies in Second Language Acquisition, 32(2), 265–302. https://doi.org/10.1017/s027263109990520

Maleki, A. (2006). The impact of the head-initial/head-final parameter on reading English as a foreign language: A hindering or a facilitating factor? The Reading Matrix, 6, 154–169.

Mao, Z., & Lee, I. (2020). Feedback scope in written corrective feedback: Analysis of empirical research in L2 contexts. Assessing Writing, 45, 1–14. https://doi.org/10.1016/j.asw.2020.100469

Mazdeh, M. M. (2013). Future. http://www.persiandee.com

McManus, K., & Marsden, E. (2019). Signatures of automaticity during practice: Explicit instruction about L1 processing routines can improve L2 grammatical processing. Applied Psycholinguistics, 40(1), 205–234. https://doi.org/10.1017/ s0142716418000553

Mohamed Salleh, R. T. A., Di Biase, B., & Ramlan, W. N. M. (2020). The acquisition of English grammar among Malay-English bilingual primary school children. GEMA Online Journal of Language Studies, 20(4), 166–185. https://doi.org/10.17576/gema-2020-2004-10

Momenzadeh, M., & Youhanaee, M. (2014). ‘Number’ and article choice: The case of Persian learners of English. Procedia-Social and Behavioral Sciences, 98, 1186–1193.

North Carolina State University. (2013). Persian in context: Grammar book to accompany units 1-8. North Carolina State University. https://cdn.chass.ncsu.edu/sites/fll.chass.ncsu.edu/persian/persian_in_context/Compiledgrammar.pdf

Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. Applied Linguistics, 10, 52–79. https://doi.org/10.1093/applin/10.1.52

Pienemann, M. (2005). Cross-linguistic aspects of processability theory (Vol. 30). John Benjamins Publishing.

Rassaei, E. (2013). Corrective feedback, learners’ perceptions, and second language development. System, 41(2), 472–483.

Rassaei, E. (2014). Scaffolded feedback, recasts, and L2 development: A sociocultural perspective. Modern Language Journal, 98(1), 417–431.

Rassaei, E. (2015). Oral corrective feedback, foreign language anxiety and L2 development. System, 49, 98–109.

Rassaei, E. (2020). The separate and combined effects of recasts and textual enhancement as two focus on form techniques on L2 development. System, 89, 1–14. https://doi.org/10.1016/j.system.2019.102193

Rassaei, E., Moinzadeh, A., & Youhanaee, M. (2012). The effect of corrective feedback on the acquisition of implicit and explicit L2 knowledge. The Journal of Language Teaching and Learning, 2(1), 59–75.

Sanavi, R. V., & Nemati, M. (2014). The effect of six different corrective feedback strategies on Iranian English language learn-
Schenck, A. (2020). Using meta-analysis of technique and timing to optimize corrective feedback for specific grammatical features. *Asian-Pacific Journal of Second and Foreign Language Education, 5*(1), 1–20. https://doi.org/10.1186/s40862-020-00097-9

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41*(2), 255–283.

Shin, M. K. (2015). The resetting of the head direction parameter. *Studies in Foreign Language Education, 18*, 17–35.

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners’ explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22*(3), 286–306. https://doi.org/10.1016/j.jslw.2013.03.011

Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners’ accuracy in using two English grammatical structures. *Language Learning, 64*(1), 103–131. https://doi.org/10.1111/lang.12029

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*(2), 263–308. https://doi.org/10.1111/j.1467-9922.2010.00562.x

Tedick, D. J., & De Gortari, B. (1998). Research on error correction and implications for classroom teaching. *ACIE Newsletter, 1*(3), 1–6.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(2), 327–369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Truscott, J. (1999). The case for “the case against grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing, 8*(2), 111–122. https://doi.org/10.1016/s1060-3743(99)80124-6

Van De Gucht, M., Braaksma, M., Rijlaarsdam, G., & Bimmel, P. (2015). Learning new grammatical structures in task-based language learning: The effects of recasts and prompts. *Modern Language Journal, 99*(2), 246–262. https://doi.org/10.1111/modl.12211

Varnosfadrani, A. D., & Basturkmen, H. (2009). The effectiveness of implicit and explicit error correction on learners’ performance. *System, 37*(1), 82–98. https://doi.org/10.1016/j.system.2008.04.004

Wang, T., & Jiang, L. (2015). The effects of written corrective feedback on Chinese EFL learners’ acquisition of English collocations. *Chinese Journal of Applied Linguistics, 38*(3), 319–338. https://doi.org/10.1515/cjal-2015-0020

Yang, M., Cooc, N., & Sheng, L. (2017). An investigation of cross-linguistic transfer between Chinese and English: A meta-analysis. *Asian-Pacific Journal of Second and Foreign Language Education, 2*(1), 1–21. https://doi.org/10.1186/s40862-017-0036-9

Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners’ acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition, 32*, 235–263. https://doi.org/10.1017/s0272263109990519

Zarei, M., Ahour, T., & Seifoori, Z. (2018). The effects of implicit, explicit, and emergent oral feedback on Iranian EFL learners’ accuracy, fluency, and attitude. *Language Horizons, 2*(2), 75–102. https://doi.org/10.22051/LGHOR.2019.26029.1122