Andrew Schenck
Fort Hays State University, Kansas, USA

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

Gleaning Insights about Input-Based and Output-Based Form-Focused Instruction through Meta-Analysis of Korean EFL Learners

A b s t r a c t

Form-Focused Instruction (FFI) has been extensively studied, yet past research often yields contradictory results. Overly simplistic examination of grammatical features, along with inadequate exploration of EFL learners at different English proficiency levels, has limited understanding of how and when to use different forms of FFI. Due to a need for further research, the present study was designed to holistically examine input and output-based forms of FFI with various grammatical features, along with Korean EFL learners at different proficiency levels. Eighteen experimental studies with Korean EFL learners were selected for metanalysis and research synthesis. The effects of type of instruction (input vs. output) were analyzed along with complexity of a target feature and learner proficiency in English. Results suggested that input-based instruction is more effective with grammatical features that involve a single phrase (e.g., basic morphology and vocabulary), while output-based instruction is more effective with more complex phrases and clauses with a syntactic component (e.g., relative clauses and conditionals). Input may activate form-meaning mapping, which is more useful for the acquisition of basic morphology or vocabulary, whereas output may help learners order constituents, thereby aiding in the acquisition of syntactic phrases and clauses. Proficiency also appears to influence the efficacy of FFI. When proficiency of participants was low and complexity of a grammatical feature was high, the influence of input tended to be stronger. When proficiency was intermediate and complexity of a grammatical feature was low, output-based instruction tended to result in larger effect sizes.

Keywords: form-focused instruction, input, output, Korean, L1, L2, EFL, grammar

Form-Focused Instruction (FFI), which refers to "any pedagogical practice aimed at drawing learners' attention to language form" (Collins & Ruivivar, 2020, p. 472), has been extensively studied (Piggott, 2019; Sun & Zhang, 2021, 2022). While FFI techniques may have a common purpose, they differ in their emphasis of either input or output. Whereas some techniques use input (written or oral) to draw attention to a grammatical feature, other techniques compel learners to produce output (speech or writing). Research suggests that both techniques may be effective (Kaivanpanah, Alavi, & Ravandpour, 2020). In a recent experimental study of 54 intermediate EFL learners in Iran, both input-based and output-based FFI treatments had a significant impact on production of vocabulary (Namaziandost, Dehkordi, & Shafiee, 2019).

Input-based instruction comes in many different forms. Input flood emphasizes a grammatical feature by placing many target features within a text, while input enhancement highlights a target feature within text through underlining or other forms of highlighting. Both techniques are minimally intrusive methods to focus attention on grammar. Regarding techniques of input enhancement like bolding, underlining, and italicizing, there is some research suggesting that it aids in language acquisition (Alanen, 1995; Lee, 2007; Rassaei, 2012, 2015; Sarkhosh et al., 2013). Input flood may also have an effect, albeit less so than input enhancement, which is more obtrusive (Rassaei, 2012). In addition to input enhancement and input flood, attention may be focused on a grammatical feature through Processing Instruction (PI). This pedagogical strategy helps learners to identify a relationship between form and meaning through structured input activities (Soruç, 2018). The technique is executed in three steps: providing information about the target linguistic form or structure, informing learners about input processing strategies that may negatively affect the target structure, and carrying out input-based activities that help the learner understand and process form during comprehension (Nassaji & Fotos, 2011, p. 24).

Like the modification of input, output may be emphasized to improve the acquisition of grammatical features. Having students produce more output, via speaking or writing, can enhance noticing of a grammatical feature. As pointed out by Swain (1998), output-based instruction can allow a language learner to test hypotheses about a grammatical feature or use "metatalk" to reflect on language use. Several studies suggest that tasks requiring output help learners to acquire grammatical features and vocabulary (Izumi, 2002; Rassaei, 2012; Shintani, 2011). In a study of 129 Persian learners, output-based instruction had the most significant impact on writing production of *so* and *such* clauses. Learners using this form of instruction outperformed peers who received both text enhancement and input enrichment (Rassaei, 2012). In a study of English relative clauses, Izumi (2002) also found that learners who engaged in output-based activities outperformed learners who received visual input.

Inconsistency of FFI Results

Although both input-based and output-based FFI have had a positive influence on acquisition of grammatical features, this efficacy is, by no means, consistent. Some research suggests that techniques such as input enhancement and input flood have little to no impact on speaking or writing (Cho, 2010; Lee, 2007; Lee & Huang, 2008; Leow et al., 2003; Rassaei, 2012). Other research suggests that the effects of PI on grammatical accuracy are limited with respect to production in speech or writing. While learners with PI treatment tend to perform better on interpretation tasks, their improvement on production tasks is often lackluster (Benati, 2005; Comer & deBenedette, 2011). Such research, along with other studies that report more significant outcomes for output-based instruction (Izumi, 2002; Rassaei, 2012, 2015), cast some doubt on the efficacy of input-based instruction.

As with input-based FFI, studies of output-based instruction often yield inconsistent results. Regarding vocabulary, Shintani (2011) found that both comprehension (input) and production (output) activities had a positive effect on production, yet the input-based tasks allowed for richer interaction that led to higher gains on a comprehension test. Benati (2001) also found that input-based interventions were more effective than output-based instruction which uses grammar rules followed by written or oral practice. Some studies of output-based FFI even suggest that the technique may be ineffectual. A study by Izumi and Bigelow (2000), for example, concluded there were no unique effects from providing an output-based task. Another study by Izumi and Izumi (2004), which examined the use of oral output, also found a nonsignificant result for output-based instruction, concluding that "re-examination of the treatment tasks revealed that the output task failed to engage learners in the syntactic processing that is necessary to trigger L2 learning, while the task for the nonoutput group appeared to promote better form-meaning mapping" (p. 587).

Past research concerning both input-based and output-based instructional emphasis of grammar has led to variable results, making interpretation of findings problematic. Without a clear ability to predict educational outcomes, teachers cannot be assured that FFI will be effective. More research is needed to provide a holistic perspective of FFI and associated influences, which may then allow educators to choose FFI techniques that are both consistent and reliable.

Timing in FFI and the Need for Further Research

Conflicting beliefs concerning the efficacy of different forms of FFI reveal a fundamental gap in existing research. Although FFI techniques may be effective, precisely how and when these techniques should be introduced to ensure consistent effectiveness is not yet known. Recently, the importance of timing in the implementation of FFI has been considered. In a study by Xu and Li (2021), FFI given before a task led to heightened efficacy for a difficult grammatical feature, whereas FFI given after a task led to heightened efficacy for an easier target feature. Such study illustrates the importance of providing FFI at an optimal time based upon grammatical feature, thereby enhancing effectiveness. While a positive step forward, it does not adequately examine issues of timing associated with different grammatical features or varied proficiency levels.

When considering the timing of FFI, characteristics of diverse grammatical features and L2 English proficiency must also be carefully considered. Research suggests that learners must be at a specific level of proficiency before a target feature can be successfully taught and acquired (Dyson, 2018; Dyson & Håkansson, 2017). According to the Processability Theory, the effectiveness of FFI is influenced by two variables, grammatical complexity and L2 English proficiency. As learners become more proficient in English as an L2, they also gain the ability to acquire (and be explicitly taught) more complex grammatical features (Pienemann & Lenzing, 2015). First, learners gain the ability to acquire intra-phrasal features that modify a single phrase. An example would be changing a present tense verb into the past tense (e.g., *eat* to *ate*; *walk* to *walked*) or making a noun into a plural (e.g., *book* to *books*). In each case, internal modification of the phrase is required. Next, learners can acquire inter-phrasal target features. This grammar requires the modification of multiple phrases on a syntactic or morphological level. An example would be inverting the auxiliary verb *do* with the subject, a noun phrase. In addition to question formation, morphological modification of a verb for the third person singular (e.g., He read*s*) reveals a cognitive connection between two phrases (the subject noun phrase and the verb). Following the acquisition of inter-phrasal features, learners who are the most proficient in English can acquire clausal features, which require the creation of larger subordinate clauses. An example of this type of grammatical feature is the conditional (e.g., *If I hadn't eaten breakfast, I would have been hungry right now*). Due to the morphosyntactic complexity of clausal features, they tend to be acquired last (Pienemann, 2005). As revealed by stages of the Processability Theory, heightened grammatical complexity may have a significant impact on acquisition, suggesting that learners need to be

at a certain English proficiency level to benefit from FFI for a specific target feature (Gholami & Zeinolabedini, 2018).

The Processability Theory appears to have implications about timing for the use of specific types of FFI. For example, there may be an optimum time to introduce input- and output-based FFI according to grammatical complexity and L2 English proficiency. Input-based instruction promotes noticing of a grammatical feature (Winke, 2013), which may prime the lexicon or serve as a scaffold when learners have had little exposure to a target feature. Output-based instruction moves learners away from "semantic, open-ended" forms of comprehension toward "complete grammatical processing needed for accurate production" (Swain, 1995, p. 128). Essentially, output pushes the learner to produce a target feature, which may be more effective when learners have already had adequate exposure to the feature and sufficient proficiency in English (Namaziandost, Dehkordi, & Shafiee, 2019). Input and output-based forms of FFI may help learners acquire a grammatical feature at an opportune time. In addition, different forms of FFI may target different forms of grammar (e.g., syntactic vs. morphological). Whereas input could provide valuable information for form-meaning mapping, output-based instruction may prompt learners to examine syntactic relationships and word order, since learners are pushed to order constituents of a phrase or sentence via this approach (Gass & Selinker, 2009). More research is needed to examine the effects of both timing and the use of different forms of FFI. This research must include a variety of target features and learners at different English proficiency levels. Currently, the correct time to introduce different forms of FFI remains unclear. Further examination of the effects of timing may explain inconsistent results of past experimentation, clarifying how and when different FFI techniques should be utilized.

Research Questions

The present meta-analysis and research synthesis was designed to investigate the effectiveness of input and output-based FFI when used with different grammatical features and learners of varied English proficiency. To aid in the investigation, the following questions were posed:

1. What types of instruction (input-based or output-based) are most effective with each type of grammatical feature (intra-phrasal, inter-phrasal, and clausal)?
2. Does the effectiveness of an instructional type (input-based or output-based) differ according to the English L2 proficiency level of a learner?

Through investigating the questions above, it was hoped that a holistic under-standing could be obtained to reform curricula or enhance automated language learning systems, thereby tailoring instruction to learner needs.

Method

The present meta-analysis synthesized several experimental studies to eval-uate the influence of grammatical complexity (intra-phasal, inter-phrasal, and clausal), type of instruction (input-based vs. output-based), and learner back-ground (English proficiency) on accuracy of production in speech or writing. Google was systematically searched by using the keywords *Korean* with various search terms for grammatical features (*plural*, *past tense*, *past regular*, *past irregular*, *passive*, *third person*, *questions*, *article*, *definite article*, *indefinite article*, *phrasal verb*, *verb particle*, *conditional*) and types of FFI treatments (*form-focused instruction*, *focus-on-form*, *focus-on forms*, *PI*, *text enhancement*, *dictogloss*, *output*, *input*, *control group*). Following the search, full texts for each study were obtained for further examination.

There are key differences between explicit knowledge and actual perfor-mance in production (implicit knowledge). Therefore, only studies that elicited responses in speech and writing were selected. To help ensure that production reflected implicit knowledge of a target feature, testing was selected that commu-nicated ideas, not rules; put pressure on learners to prevent conscious correction of language errors; focused on meaning not form; and avoided use of metalan-guage (Ellis, 2009). In order to be included within the present meta-analysis, each experimental study needed to have:

1. An input-based or output-based treatment (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. Participants with Korean as their L1.

Information about proficiency was often variable and inconsistent. Studies that did include this variable were placed in basic categories such as beginner, intermediate, and advanced for further examination in this study.

Originally 59 experimental studies with South Korean EFL learners were located. From this group, only 18 studies met the criteria for inclusion. Many studies lacked adequate assessment of productive, implicit knowledge, leading to exclusion (See Appendix B for more information on assessments used in the selected studies). In addition to problems with assessment of production, some studies lacked sufficient knowledge to understand the methodology or length

of treatment. Thus, they were excluded. Totally, the present meta-analysis contained 18 studies for analysis. In total, there were 47 treatment groups.

Grammatical Feature Type

After studies were compiled, results from each study were organized and evaluated according to grammatical feature. Types of grammatical feature were grouped for comparison. As revealed by the Processability Theory, grammatical complexity could vary based on whether a target feature is intra-phrasal (e.g., a verb or noun and associated morphological features like the regular past -*ed* or plural -*s*), inter-phrasal (e.g., question inversion or phrasal verbs), or clausal (e.g., relative clauses or conditionals) (Pienemann & Lenzing, 2015). Grammatical features from the selected studies were taken from Appendix A and separated as in Table 1.

Table 1

Separation of Grammatical Features Based on Processability Theory

	Target Features (Korean Studies)
Intra-phrasal	Single adjectives, verbs, and nouns (2 Studies) Present Perfect (4 Studies) Collocations (despair, earnestly, etc.) (1 Study) Verb tenses (1 Study) Participial adjectives (e.g., *boring/bored*) (1 Study)
Inter-phrasal	Comparative adjectives and *than* (smarter than X) (1 Study) Verb + Pronoun + to + V (I want her to visit) (1 Study)
Clausal	Conditional (5 Studies) That-Trace Filter (1 Study) Relative Clauses (1 Study)

In total, 47 treatment groups were obtained from the selected studies. Because nine of the control groups contained collocations with more than one grammatical feature from a different complexity level, these treatment groups were excluded from analysis of grammar type, leaving a remainder of 38 treatment groups.

Proficiency Level

Studies chosen for meta-analysis could include different proficiency designations such as low beginner, beginner, high beginner, intermediate low, intermediate, or high intermediate (See Appendix A). However, this variation was not systematic or complete. As an example, the beginner level included only students who were evaluated as "low English proficiency" by using TOEIC scores that ranged from 250 to 320. Due to variability in how students were assessed as either high or low, level designations were collated into three general groups: beginner, intermediate, and advanced. Of the 18 studies with Korean participants, only 14 had information for designation of proficiency level (8 beginner treatment groups and 31 intermediate treatment groups). Information provided (or not provided) concerning proficiency levels confirms assertions by Liu and Brown (2015), who contend that methodological inconsistencies have limited our understanding of ESL or EFL instruction.

While studies selected for meta-analysis did generally include some assessment of levels, they should be interpreted with caution. Because different instruments were used to assess proficiency, discrepancies between assessment of proficiency may be expected. Ultimately, assignment of proficiency levels in selected studies reflects the researcher's judgment, meaning that this variable cannot be interpreted as an empirical variable based on a standardized assessment. It must be considered a construct of the researcher, like variables obtained from survey data. Classifications of proficiency level may reveal some trends that can be confirmed later through follow-up experimental research with more standardized instruments.

Input-Based Output-Based Definitions

Studies designed to evaluate the efficacy of either input or output were selected and separated based on instructional type (See Appendix B for more information about treatments). Any studies that had a major emphasis on both input and output in the FFI treatment were excluded from analysis. Whereas treatments primarily designed to emphasize the impact of input (e.g., input flood, IE, and PI) were assigned to the input category, tasks that emphasized output (e.g., text reconstruction or dictogloss) were assigned to the output category. The input vs. output distinction was used to analyze differences in effect sizes, along with variables such as grammatical complexity, learner proficiency, and

learner L1. For input, there were a total of 30 treatment groups; for output, there were a total of 17 groups.

Overall, output-based treatments included a variety of both written and spoken tasks. Among these tasks, the dictogloss encouraged both written and spoken output. As a text is being reconstructed, learners may use metatalk to discuss the target feature. Other studies of output-based assessments like that of Kim (2014) utilized images to elicit verbal responses about a target feature. With the exception of studies that used only the dictogloss, there was little standardization of the techniques used to elicit output. While forms of production did vary, Swain (1998) points out that all production tasks give learners the ability to use and test hypotheses about a target feature. Some output treatments did include a degree of explicit information or guidance to conduct the activity, which was a type of input. In each treatment, however, emphasis was placed on producing output, rather than providing input. Studies that used a dictogloss, for example, required input before the story was reconstructed. However, the main goal of the activity was output, as reflected by procedures that included notetaking, meta-talk, and story construction. Although some input may have been provided with output treatment groups, the main goal of these groups was to produce either an oral or written product. Therefore, a clear emphasis was placed on production, rather than input, allowing learners time to test hypotheses as envisioned by Swain (1998). Some studies provided an emphasis on both input and output in the same treatment groups. These treatment groups were excluded from analysis. Control groups with no treatment were also excluded from analysis.

L1 Influence

Research suggests that L1 transfer does indeed have an impact on acquisition, thereby impacting the effectiveness of FFI (Luk & Shirai, 2009; Maleki, 2006; Shin, 2015; Yang et al., 2017). Therefore, EFL learners with the same native language (Korean) were studied. Examination of learners with the same L1 helped to ensure that differences in experimental results were related to grammatical feature type and L2 English proficiency, rather than differences due to L1 transfer. Grammatical features targeted in experimental studies emphasized English grammatical features that were very different from Korean. Studies emphasized grammatical features like relative clauses or conditionals, which use a head final parameter (opposite from the predominant head initial parameter in English). Other studies examined verb tenses. Korean verb tenses for the simple past resemble English in some ways, yet the present perfect and

present perfect progressive tenses lack equivalent morphosyntactic structures in Korean. Instead, Korean uses a variety of morphological verb endings and adverbials to express aspect. Other features such as participials (e.g., *I'm bored* vs. *The book is boring*) and verb/pronoun combinations (e.g., *ask me to*) also lack equivalent structures in Korean. Overall, English grammar emphasized within FFI studies was significantly different from grammatical features found in the Korean L1 (See Appendix A for more information on target features used for each experimental study).

Procedure

To compare results from individual studies, effect sizes needed to be calculated for each study. Whereas *p* values reveal whether results are significant (not a result of chance), effect sizes determine the magnitude of a difference between groups (Sullivan & Feinn, 2012). Since significant *p* values may not actually reveal a large effect (e.g., large numbers of participants or amounts of data may reveal only a small significant difference), effect size is needed to understand how effectual a treatment is. Calculation of effect size also provides a consistent way to compare different studies, since the calculations are standardized measures.

Cohen's *d* was selected since it is widely used to compare two means within meta-analysis studies and provides "an important complement to traditional null hypothesis statistical significance testing" (Diener, 2010, p. 1). It was also used in the formative meta-analysis of FFI conducted by Spada and Tomita (2010), which analyzed effects of explicit and implicit instruction on the acquisition of simple and complex grammatical features in English. In the current study, effect size was calculated by inserting pretest scores (M2), posttest scores (M1), and associated standard deviations (SD2 and SD1) into the Cohen's *d* formula for effect size (Spada & Tomita, 2010, p. 307):

$$d = [M1 - M2] / [SQRT[(SD1SD1 + SD2SD2)/2]$$

After effect sizes were calculated for each treatment group, results were collated based upon the variables, allowing for further analysis. For grammatical complexity, effect sizes were collated based upon whether an intra-phrasal, inter-phrasal, or clausal feature was emphasized. For proficiency level, effect sizes were collated according to researcher designations of beginner or intermediate (there were no advanced). Average effect sizes for both grammatical complexity and proficiency level were then subdivided based upon instructional type (input or output) for further analysis.

Results and Discussion

Grammatical Complexity and the Efficacy of FFI

Research of grammatical features revealed interesting differences between the use of input-based and output-based instruction. Table 2 revealed that input-based instruction was more effective when less complex intra-phrasal features were emphasized, while output-based instruction was more effective for inter-phrasal and clausal features. As for intra-phrasal features, input-based instruction had a larger average impact of .42 over its output-based counterpart. Because prior research of meta-analysis suggests that a small effect size is $d > 0.2$, a medium effect size is $d > 0.5$, and a large effect size is $d < 0.8$ (Rice & Harris, 2005), this value would represent a small but substantial difference. Inter-phrasal features revealed a much higher value for output-based instruction, which represented a difference in effect size of .67, which would be a medium effect. Although significantly different, this value should be interpreted with caution. The output-based score is determined through only one study by Hwang (2018), which examined comparative adjectives with *than*. Concerning clausal features, output-based instruction was more effective by .09.

Table 2

Average Effect Size by Grammatical Complexity and Type of Instruction

	Input-Based vs. Output-Based	Mean	*N*	Std. Deviation
Intra-phrasal	Input	1.9548	10	1.08530
	Output	1.5345	8	.70399
Inter-phrasal	Input	1.4015	6	.63283
	Output	2.0699	1	.
Clausal	Input	1.2421	7	1.11637
	Output	1.3345	6	1.10216

Intra-phrasal features tended to benefit more from input-based instruction. Experimental studies of vocabulary and collocations by Lee (2002) and Lee (2003) both show higher effect sizes for input over output. Experimental studies that examine verb tenses also show a greater impact for input (Yang, 2004, 2008; Hwang, 2018). Findings may be explained by the characteristics of target features, which have a large number of lexical elements that require form-meaning mapping. Input may provide the semantic and morphological information necessary to prime the learner, thereby aiding in the acquisition of form-

meaning mappings. This perspective may also explain the results of a study by Kim and Nam (2017), which revealed that emphasis of inter-phrasal collocations such as "shake a leg" or "hit the sack" via input-based instruction yielded very high effect sizes that averaged 3.25.

Input may provide scaffolding for morphological and lexical grammatical features that require form-meaning mapping early in the acquisition process. As the complexity of syntax increases, output-based instruction may be more effective, pushing learners to arrange syntactic elements into the correct order. Concerning clausal features, output-based instruction was more effective for both inter-phrasal and clausal features. This finding may reflect an effect of output-based instruction on more syntactically complex features that require ordering of constituents. Experimental studies which had both input and output-based treatments for conditionals (Song, 2007; Shin, 2011) had higher effect sizes for output. A study of relative clauses by Hwang (2018) likewise shows a larger effect size for output over input. When producing output, the learner is forced to attend to word order, thereby producing a correct oral or written utterance. Output-based activities may activate cognitive mechanisms that deal with syntax.

Input or output-based instruction may emphasize different characteristics of grammar. Since the difficulty posed by lexical variation or syntactic complexity is also partially determined by L2 proficiency, characteristics of the learner must be further studied to clarify the impact of input- and output-based FFI.

English Proficiency and the Efficacy of FFI

Separation of effect size based upon proficiency suggests that output-based instruction is more effective at the beginner level, while input-based instruction is more effective at the intermediate level (Table 3).

Table 3

Mean Effect Sizes for Input and Output-Based Instruction at Two Proficiency Levels

	Input-Based vs. Output-Based	Mean	N	Std. Deviation
Beginner	Input	2.1042	4	.82168
	Output	2.3800	4	.97856
Intermediate	Input	1.8029	21	1.25458
	Output	1.0781	10	.38060

At the beginner level, the difference is .28, a small magnitude. Although this finding appears to suggest that input-based instruction is more beneficial for beginner learners, the findings should be interpreted with caution. Results for the beginner levels were based upon just one study by Hwang (2018), which analyzed verb tenses (intra-phrasal), comparative adjectives with *than* (inter-phrasal), and relative clauses (clausal). The study revealed that input-based instruction was more effective for both verb tenses (difference in effect size of .47) and comparatives (difference in effect size of .18). Output-based instruction was more effective for relative clauses (difference in effect size of 2.28). Interestingly, the effect of output-based instruction systematically increases, and the effect of input-based instruction decreases as the grammatical feature becomes more complex. Thus, the study lends further for the idea that input-based instruction is effective for intra-phrasal morphology, whereas output-based instruction is more effective for inter-phrasal and clausal grammatical features.

At the intermediate proficiency level, the difference is substantial (a difference of .72), which suggests a medium effect. The results were heavily influenced by studies like those of Yang (2004, 2008), who emphasized the present perfect progressive, and Kim and Nam (2017), who emphasized inter-phrasal lexical collocations. All of these studies had higher effect sizes for input, lending further support for the idea that input is more effective for form-meaning mapping of morphological and lexical features.

In studies by Kim (2002) and Yeo (2002), which emphasized participial adjectives (*boring/bored*), learners performed better when output-based instruction was used (differences in effect sizes of 1.00 and .52, respectively). Unlike the more complex morphosyntax of present perfect progressive tenses emphasized by Yang (2004, 2008) or inter-phrasal collocations emphasized by Kim and Nam (2017), participial phrases are simple, binary morphemes. Perhaps low complexity of the target feature, along with intermediate English L2 proficiency explains the findings. Output may be more effective with simplistic morphological features when English L2 proficiency is high enough. In the study by Kim (2002), an inter-phrasal feature (verb + *me* + *to*) was also included in the same treatment and assessment of participial phrases. Inclusion of a syntactic feature may explain why the output-based instruction was nearly twice as effective than that of Yeo (2002), who examined only participial adjectives. While a potentially valid interpretation of research results, it is important to note that the studies by Yeo (2002) and Yang (2004) did not include information about learner proficiency. Further research is needed to ensure that proficiency level influences FFI outcomes for target features such as participial adjectives or present perfect tenses.

Conclusion

The present study was designed to facilitate understanding of how input- and output-based FFI can be utilized in a more timely and effective way. Results appear to suggest that efficacy of either input or output-based instruction is dependent on characteristics of the target feature. Simplistic intra-phrasal features (morphological and lexical features) may benefit more from input. Input-based instruction allows learners to consider form-meaning mappings, without a need to attend to issues of word order. More complex inter-phrasal and clausal features may benefit more from output-based instruction. In contrast to input-based FFI, output-based instruction forces the learner to address issues related to word order, which may explain why inter-phrasal and clausal syntactic features (e.g., subject/verb inversion, relative clauses, or conditionals) had larger effect sizes for this instructional style.

There may also be a relationship between L2 English proficiency and the efficacy of different types of FFI. While collation of effect sizes appears to suggest that beginner levels benefit more from output and intermediate levels benefit more from input, systematic exploration of individual studies appears to suggest the opposite. Using average effect sizes for proficiency does not include consideration of the challenges posed by complexity of a target feature. Results from individual studies provide some evidence that simplicity of a target feature and intermediate proficiency make output-based instruction more effective. At lower proficiency levels, input appears to be more effective when complexity of a grammatical feature is high. Such findings reveal a close link between grammatical feature and proficiency level. Both variables may need to be considered before the effectiveness of different FFI types can be accurately assessed. Although conclusions about proficiency are interesting, they must be interpreted with caution. Assessment of English proficiency was not consistent within past studies, which may have influenced the findings.

Results of the present study have implications for both instruction and the development of educational technology. Input may need to be modified (or remain unmodified) based upon the target feature chosen for emphasis. Whereas modifying input can promote acquisition of intra-phrasal features related to verb tense or noun forms, it may be less effective for more complex inter-phrasal and clausal features. Rather than input-based FFI, enhanced output-based tasks may be used to promote the acquisition of more complex syntactic features such as question inversion or relative clauses. While a target feature partially dictates whether input- or output-based FFI should be used, a learner's proficiency level should also be considered. Learners must be at a proficiency level which allows for the acquisition of a target feature. As proficiency rises, cognitive resources or abilities may also increase, meaning that different tasks (input vs.

output) could be effective at different times. In the future, language instruction and educational technology will need to diversify FFI, carefully controlling administration of techniques based upon target feature and proficiency level. In this way, English instruction may be tailored to the unique needs of each individual learner.

While some very insightful information was gleaned from existing studies of FFI, there were key limitations that need to be addressed in future research. The number of grammatical features emphasized in these studies was limited, as was the variety of L2 proficiency levels examined in the studies. Clearly, more research is needed. Experimental studies conducted to date often examine the same grammatical features or use learners at similar proficiency levels (e.g., intermediate). As a result, the impact of different target features and timing (introduction at the right level of L2 proficiency) is unknown. We must fully examine the range of grammatical features presented at a variety of proficiency levels, thereby leading to a concrete understanding of how and when to use FFI.

References

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259–302). University of Hawaii Press.

Benati, A. (2001). A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense. *Language Teaching Research, 5*(2), 95–127. https://doi.org/10.1177/136216880100500202

Benati, A. (2005). The effects of processing instruction, traditional instruction and meaning—Output instruction on the acquisition of the English past simple tense. *Language Teaching Research, 9*(1), 67–93. https://doi.org/10.1191/1362168805lr154oa

Cho, M. Y. (2010). The effects of input enhancement and written recall on noticing and acquisition. *Innovation in Language Learning and Teaching, 4*(1), 71–87. https://doi.org/10.1080/17501220903388900

*Cho, Y. (2011). Effects of different types of focus-on-form instruction and learner readiness on the learning of perfective modals. *Foreign Languages Education, 18*, 1–27.

Collins, L., & Ruivivar, J. (2020). Form-Focused Instruction. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 472–502). John Wiley and Sons. https://doi.org/10.1002/9781405198431.wbeal0432.pub2

Comer, W. J., & deBenedette, L. (2011). Processing instruction and Russian: Further evidence is in. *Foreign Language Annals, 44*(4), 646–673. https://doi.org/10.1111/j.1944-9720.2011.01155.x

Diener, M. J. (2010). Cohen's *d*. In *The Corsini encyclopedia of psychology*. https://doi.org/10.1002/9780470479216.corpsy0200

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

Gass, S. M., & Selinker, L. (2009). *Second language acquisition: An introductory course* (3rd ed.). Routledge.

Gholami, J., & Zeinolabedini, M. (2018). Learnability and teachability hypothesis. In J. I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1–7). John Wiley & Sons.

*Hwang, H. J. (2018). A study on the Korean EFL learners' grammatical knowledge development under input-enhanced FFI and output-enhanced FFI conditions. *Journal of Digital Convergence*, *16*(5), 435–443. https://doi.org/10.14400/JDC.2018.16.5.435

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*(4), 541–577. https://doi.org/10.1017/S0272263102004023

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition?. *TESOL Quarterly, 34*(2), 239–278. https://doi.org/10.2307/3587952

Izumi, Y., & Izumi, S. (2004). Investigating the effects of oral output on the learning of relative clauses in English: Issues in the psycholinguistic requirements for effective output tasks. *Canadian Modern Language Review, 60*(5), 587–609. https://doi.org/10.3138/cmlr.60.5.587

*Jeong, H., & Lee, H. (2018). The effects of two types of written output tasks on noticing and learning of vocabulary. *Korean Journal of Applied Linguistics, 34*(4), 75–102. https://doi.org/10.17154/kjal.2018.12.34.4.75

Kaivanpanah, S., Alavi, S. M., & Ravandpour, A. (2020). The effect of input-based and output-based tasks with different and identical involvement loads on Iranian EFL learners' incidental vocabulary learning. *Cogent Psychology, 7*(1), 1–17. https://doi.org/10.1080/23311908.2020.1731223

*Kang, D. (2011). Noticing and L2 learning in input-based focus-on-form instruction. *Foreign Languages Education, 18*(2), 29–49. UCI: G704-000287.2011.18.2.017

*Kang, D. H. (2003). Focus on form instructions and L2 learners' instructional preferences. *Foreign Languages Education*, *10*(1), 57–82. UCI: G704-000287.2003.10.1.007

*Kang, N. (2009). Effects of form-focused instructions on the learning of English verb complementation by Korean EFL learners. *English Teaching, 64*(1), 3–25. https://doi.org/10.15858/engtea.64.1.200903.3

*Kim, B. (2002). The role of focus on form in the EFL classroom setting: Looking into recast and CR tasks. *English Teaching, 57*(4), 267–296. UCI: G704-000189.2002.57.4.022

*Kim, J. E. (2014). Timing of form-focused instruction and development of implicit vs. explicit knowledge. *English Teaching, 69*(2), 123–147.

*Kim, J. E., & Nam, H. (2017). The pedagogical relevance of processing instruction in second language idiom acquisition. *International Review of Applied Linguistics in Language Teaching, 55*(2), 93–132. https://doi.org/10.1515/iral-2015-0027

*Lee, H. (2002). Communicative output as a mode of focus on form. *English Teaching, 57*(3), 171–192. UCI: G704-000189.2002.57.3.006

*Lee, H. (2003). The effects of production and comprehension for focus on form and second language acquisition. *Korean Journal of Applied Linguistics, 19*(2), 41–68. UCI: G704-000491.2003.19.2.010

Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning, 57*(1), 87–118. https://doi.org/10.1111/j.1467-9922.2007.00400.x

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*(3), 307–331. https://doi.org/10.1017/S0272263108080479

Leow, R. P., Egi, T., Nuevo, A. M., & Tsai, Y. C. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning, 13*(2), 1–16.

*Lim, J. (2007). Input enhancement in the EFL learning of present perfect and the lexical aspect. *Modern English Education, 8*(3), 38–60. UCI: G704-001860.2007.8.3.005

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 's. *Language Learning, 59*(4), 721–754. https://doi.org/10.1111/j.1467-9922.2009.00524.x

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.

Namaziandost, E., Dehkordi, E. S., & Shafiee, S. (2019). Comparing the effectiveness of input-based and output-based activities on productive knowledge of vocabulary among pre-intermediate EFL learners. *Asian-Pacific Journal of Second and Foreign Language Education, 4*, 1–14. https://doi.org/10.1186/s40862-019-0065-7

Nassaji, H., & Fotos, S. S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. Routledge.

Pienemann, M. (2005). *Cross-linguistic aspects of Processability Theory* (Vol. 30). John Benjamins Publishing.

Pienemann, M., & Lenzing, A. (2015). Processability Theory (2nd ed.). In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 159–179). Routledge. https://doi.org/10.4324/9780203628942

Piggott, L. (2019). *First meaning then form: A longitudinal study on the effects of delaying and reducing explicit form-focused instruction for young adolescent EFL learners*. Doctoral dissertation, Utrecht University, The Netherlands. LOT Dissertation Series, number 545. LOT.

Rassaei, E. (2012). The effects of input-based and output-based instruction on L2 development. *TESL-EJ, 16*(3), 1–26. https://tesl-ej.org/wordpress/issues/volume16/ej63/ej63a2/

Rassaei, E. (2015). Effects of textual enhancement and input enrichment on L2 development. *TESOL Journal, 6*(2), 281–301. https://doi.org/10.1002/tesj.149

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior, 29*(5), 615–620. https://doi.org/10.1007/s10979-005-6832-7

Sarkhosh, M., Taghipour, B., & Sarkhosh, H. (2013). Differential effect of different textual enhancement formats on intake. *Procedia-Social and Behavioral Sciences, 70*, 544–559. https://doi.org/10.1016/j.sbspro.2013.01.092

*Shin, M. K. (2011). Effects of output tasks on Korean EFL learners' noticing and learning of English grammar. *Foreign Language Education, 18*(2), 127–163. UCI: G704-000287.2011.18.2.014

Shin, M. K. (2015). The resetting of the head direction parameter. *Studies in Foreign Language Education, 18*, 17–35. http://hdl.handle.net/10371/95840

Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research, 15*(2), 137–158. https://doi.org/10.1177/1362168810388692

*Song, M. J. (2007). Getting learners' attention: Typographical input enhancement, output, and their combined effects. *English Teaching, 62*(2), 193–215. https://doi.org/10.15858/engtea.62.2.200706.193

*Song, M. J., & Suh, B. R. (2008). The effects of output task types on noticing and learning of the English past counterfactual conditional. *System, 36*(2), 295–312. https://doi.org/10.1016/j.system.2007.09.006

Soruç, A. (2018). The mediating role of explicit information in processing instruction and production-based instruction on second language morphological development. *Educational Sciences: Theory & Practice, 18*(3), 279–302. https://dergipark.org.tr/en/pub/buje/issue/55539/760450

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

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the *P* value is not enough. *Journal of Graduate Medical Education, 4*(3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1

Sun, Q., & Zhang, L. J. (2021). A sociocultural perspective on English-as-a-foreign-language (EFL) teachers' cognitions about form-focused instruction. *Frontiers in Psychology, 12*, 1–14. https://doi.org/10.3389/fpsyg.2021.710736

Sun, Q., & Zhang, L. J. (2022). Understanding novice and experienced teachers' cognitions and practices for sustainable teacher development: the case of form-focused instruction in English language teaching. *Sustainability, 14*(8), 1–17. https://doi.org/10.3390/su14084711

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125–144). Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64–81). Cambridge University Press.

Winke, P. M. (2013). The effects of input enhancement on grammar learning and comprehension: A modified replication of Lee (2007) with eye-movement data. *Studies in Second Language Acquisition, 35*(2), 323–352. https://doi.org/10.1017/S0272263112000903

Xu, J., & Li, C. (2021). The effects of the timing of form-focused instruction on EFL learners' learning of difficult and easy grammatical features: A comparative study. *System, 101*. https://doi.org/10.1016/j.system.2021.102612

*Yang, M. (2004). A comparison of the efficacy between input-based instruction and output-based instruction in focus on form. *English Teaching, 59*(2), 145–164. UCI: G704-000189.2004.59.2.012

*Yang, M. (2008). The effects of task difficulty on learners' attention to meaning and form during focus-on-form instruction. *Foreign Languages Education, 15*(3), 27–52. UCI: G704-000287.2008.15.3.006

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*(15), 1–21. https://doi.org/10.1186/s40862-017-0036-9

*Yeo, K. (2002). The effects of dictogloss: A technique of focus on form. *English Teaching, 57*(1), 149–167. UCI: G704-000189.2002.57.1.011

Declarations:

Studies of Input/Output Using Learners with Korean L1

Authors	Duration	Learner Proficiency (Number)	Grammar Feature
Kang (2011)	3 hours in one week	TOEIC 420 to 930 (Average 675) (n = 15)	conditional
Kang (2003)	12 hours / 6 hours for two weeks	no info (n = 134)	conditional
Song and Suh (2008)	3 sessions over a month period	Intermediate (n = 52)	conditional
Yang (2008)	2-hour treatment	approximately 140 to 185 on reading TOEIC (n = 70)	present perfect progressive
Kim (2014)	3 sessions of treatment over 3 days	Intermediate level self-rating (n = 42)	*Who do you think that married Sarah last year?* / that-trace filter
Kang (2009)	eight 25-minute form focused treatment over 4 weeks	Intermediate (n = 150)	Verb + NP + to infinitive / ask me to
Song (2007)	3 sessions of about 30 minutes over a 3-week period	intermediate TEPS 501 to 700 (average 600.5) (n = 140)	conditional
Yang (2004)	2-hour class	no info (n = 70)	present perfect progressive
Kim (2002)	eight 25-minute focus on form treatments over four weeks	low intermediate (n = 72)	participial adjectives (*boring*) and Verb + Pronoun + to + V (I want her to visit my place)
Yeo (2002)	20 to 30 minutes twice (once each week for two weeks)	no info (n = 90)	participial adjectives

Appendix A continued

Cho (2011)	25 minutes in one session	no info (*n* = 117)	present perfect (for regret)
Lee (2003)	One session (reading session 5 mins long and interaction session 8 to 10 mins)	low to mid intermediate (*n* = 68)	words like *eyesight* and *adequate*
Jeong and Lee (2018)	Treatment session 1 (30 min) + Treatment session 2 (after 5 min. break / 30 min)	high intermediate (*n* = 66)	two adjectives and five nouns
Lim (2007)	two 50-minute classes a week for 3 weeks (total of 300 minutes) - 6 sessions over a three-week period	Intermediate (*n* = 62)	Present Perfect
Shin (2011)	Over 2 weeks	both low and high mixed (Based on high school midterm exam) (*n* = 90)	Conditional
Hwang (2018)	One class period	250 to 320 TOEIC (low English Proficiency) (*n* = 122)	Verb tenses and relative clauses
Lee (2002)	One class period	low intermediate (*n* = 124)	Collocations such as: I had better do/at an American college / smart enough to / I had better do / turned to / in despair
Kim and Nam (2017)	All on same day / Pretest – Treatment of around 20 to 30 minutes	high intermediate or low advanced by self-report (*n* = 100)	10 Idiomatic expressions like "shake a leg" and "hit the sack"

Appendix B

Number of Treatments, Treatment Types, and Assessment

Korean Studies	
Authors Number of Groups – Treatment Type – Effect Size	Assessment
Kang (2011) 1 – Input (Low proficiency input enhancement and conscious- ness raising) (d = 3.3120) 1 – Input (High proficiency input enhancement and conscious- ness raising) (d = .7846)	Dictogloss (written)
Kang (2003) 1 – Output (All output and production activities) (d = 1.6461) 1 – (Processing Instruction) (d = 1.5110)	Sentence Production Task (written)
Song and Suh (2008) 1 – Output (Reconstruction task) (d = .7611) 1 – Output (Picture-cued writing) (d = 1.0041)	Contextual sentence com- pletion task (production) / TLU Analysis Used (written)
Yang (2008) 1 – Input (Textual Enhancement) (d = 3.1854) 1 – Output (Dictogloss) (d = 2.6561)	Production test where Korean asked for English response / ten minutes / (written)
Kim (2014) 1 – Input (60 target sentences with images and rule explana- tion in Korean) (d = 1.1846) 1 – Input (60 target sentences with images and rule explana- tion in Korean) (d = .7914) 1 – Input (Input Enhancement - 60 target sentences with images) (d = .7770)	Oral imitation test
Kang (2009) 1 – Input (negative feedback plus one session of rule presen- tation) (d = 2.0896) 1 – Input (Input Enhancement) (d = 1.3206)	Picture Description Task (written)
Song (2007) 1 – Input (input enhancement) (d = −.0654) 1 – Output (picture cued writing) (d = 1.0042)	Production test about 30 minutes / explained in Korean with prompt (written)
Yang (2004) 1 – Input (in-depth reading, comprehension questions, and feedback - student metatalk but not implicit treatment given) (d = 3.1854) 1 – Output (a modified dictogloss) (d = 2.6140)	Sentence production test / translate Korean sentences into English (written)

Appendix B continued

Kim (2002)
1 – Output (communicative tasks designed to elicit target form)
(*d* = 1.2692)
1 – Input (Consciousness Raising - Rules of target forms /
comprehension and checkup questions) (*d* = .2630)

Elicited oral production test and picture description

Yeo (2002)
1 – Input (Input
Enhancement) (*d* = .6998)
1 – Output (Dictogloss) (*d* = 1.2229)

Open ended oral questions

Cho (2011)
1 – Input (input flood and input enhancement combined)
(*d* = 1.0470)
1 – Output (Image and sentence completion) (*d* = 1.5110)

Fill-in-the-blank and open-ended questions (written)

Lee (2003)
2 – Input (input with target feature) (*d* = 1.1159)
2 – Output (production activities) (*d* = 1.0066)

Uses some comprehension measures but has a major production component (written)

Jeong and Lee
(2018)
1 – Input (Input
Enhancement with pictures – body part circled) (*d* = 3.7538)

Fill in the blank production test about parts of the body (written)

Lim (2007)
1 – Input (Input with visual) (*d* = 1.4194)

Blanks of one to up to three words to assess productive knowledge / 14 verbs from three aspectual categories and 20 distractors (written)

Shin (2011)
1 – Output (Guided Essay) (*d* = 1.0495)
1 – Output (Reconstruction task) (*d* = .6293)
1 – Input (input with target feature) (*d* = .2237)

Picture-Cued Production test and fill-in-the-blank (contextualized pictures)— 8 targets and 4 distractors (written)

Hwang (2018)
3 – Input (Input Enhancement—narratives with bolded text)
(*d* = 1.7438)
3 – Output (Audio-recorded narratives with dictogloss) (*d* = 2.2880)

Production with fill-in-the-blank and open-ended questions (written)

Lee (2002)
2 – Output (read 146-word story out loud and worked
in groups to retell story) (*d* = 1.5252)
2 – Input (with picture prompts) (*d* = 1.7565)

Production through translation and sentence completion (5 minutes given) (written)

Kim and Nam
(2017)
3 – Input (explicit information provided) (*d* = 3.5602)
2 – Input (with target feature) (*d* = 2.7907)

Production task with ten short contexts and blank for idiomatic expression (written)