Instruction on the English Tense System through Systemic Theoretical Instruction and Cognitive Grammar: Impacts on Students’ Grammatical Production

Chi Wui NG
The Chinese University of Hong Kong

Abstract

Traditional grammar pedagogy conceptualizing grammar as rules of thumb fails to provide learners with comprehensive, accurate, or systematic knowledge on the system. The present classroom-based study aims at investigating application of an alternative pedagogical grammar method integrating the pedagogical framework of systemic theoretical instruction and the linguistic framework of cognitive grammar to instruction on the English tense system. Four instructional sessions were conducted in an English language classroom in a Hong Kong secondary school, and impacts of the pedagogy on students’ grammatical production were examined by means of tests and focus group interviews. The pedagogical approach was discovered to exert limited impacts on students’ controlled and free grammatical production statistically, yet the effect size of the pedagogy on students’ grammatical production was large. Several factors ought to be taken into consideration in application of the pedagogy to second language grammar instruction.

Attributed to conclusive evidence for higher pedagogical efficacy of explicit grammar instruction when compared to that of implicit grammar instruction (Ellis, 2002a; Norris & Ortega, 2000), an indispensable role of explicit grammar instruction in second language (henceforth L2) learning has been corroborated (Celce-Murcia, 2015; Ellis, 2002b, 2006; Housen & Pierrard, 2006). Drawing upon Halliday’s (1977) Systemic Functional Linguistics (SFL), Larsen-Freeman (2003) has incorporated dimensions of form, meaning, and use into her pedagogical framework for grammar description; students are expected to be equipped with a skill of grammaring, which denotes ability to capitalize upon language forms to convey meanings in appropriate social contexts.

In light of the advent of such a conceptualization of grammar, it is not uncommon to see a number of grammar pedagogy, videlicet processing instruction, collaborative output tasks, and discourse-based grammar instruction, exploited by L2 educators for grammar instruction. Targeting on learners’ language processing pitfalls resulting from their input processing strategies identified, processing instruction aims at assisting learners in processing target forms correctly for meaning by means of specifically designed structured input activities (Van Patten, 1996, 2004). Grounded upon Swain’s (1985, 1995) Output Hypothesis, collaborative output tasks, such as dictogloss tasks, are meaning-based communicative tasks intended to promote learning of target structures through provision of primary language data and elicitation of pushed output from learners (Wajnryb, 1990). Merging discourse analysis and grammar instruction, form-focused discourse hooks learners’ attention to grammar usage in authentic language at a discourse level (Celce-Murcia & Olshtain, 2001). For all their concurrence with Long’s (1985, 1991) focus-on-form (FoF) approach to grammar instruction, the three instructional methods conceptualize grammar as rules of thumb dissociating syntax from semantics and segregating language use from
human cognition. They are inadequate in providing learners with comprehensive, accurate, or systematic knowledge on the tense system, let alone genuinely equipping them with grammering skills (Negueruela, 2008; Tyler, 2012). Pedagogical alternatives providing learners with precise conceptual understanding of grammar are thereby warranted; this motivates proposal for the pedagogical grammar method investigated in the present study.

Capable of advancing students’ conceptual development via introduction of semantic concepts underlying target language structures, a combination of systemic theoretical instruction, which is grounded upon a learning theory named sociocultural theory, and cognitive grammar, which is a hyponym of a usage-based linguistic theory named cognitive linguistics, offers an alternative for instruction complementing the inadequacy of the three aforementioned pedagogical approaches. All the same, empirical support for such an alternative appears frightfully limited.

The present study thereby aims at investigating application of the pedagogy to instruction on the English tense system in the context of a Hong Kong secondary school. In accordance with Purpura’s (2004) model of grammatical constructs, grammatical performance is manifestation of grammatical knowledge. Grammatical knowledge, which possesses both receptive and productive respects, can be measured via grammatical performance (Purpura, 2004). Being an abstract theoretical construct, hardly can grammatical knowledge, which comprises grammatical comprehension and grammatical production, be directly assessed and ought to be measured via grammatical performance, which is a hyponym of language performance (ibid.). Grammatical production, which can further be divided into controlled and free grammatical production, constitutes the focus of the present paper. In particular, the current study is intended to address the following research questions:

1. To what extent does instruction integrating STI with CG ameliorate students’ controlled grammatical production of English present and past tenses?
2. To what extent does instruction integrating STI with CG ameliorate students’ free grammatical production of English present and past tenses?

**Systemic Theoretical Instruction**

Systemic Theoretical Instruction (STI), which is a pedagogical model developed by Gal’perin (1969, 1979, 1989, 1992), provides the pedagogical framework for the pedagogical grammar method under investigation in the study. STI possesses four tenets: concepts as minimal units of instruction, materialization of concepts, verbalization of concepts, and interconnection amongst categories of meaning (Negueruela, 2003, 2008). Pedagogical procedures of STI proposed by Gal’perin (1969) and outlined by Haenen (1996) are delineated below.

STI commences with the motivational stage, where learners are motivated to learn target language structures via provision of primary language data comprising instances of target structures (Haenen, 1996). Following the motivational stage are orienting and material(ized) stages, where semantic and pragmatic concepts underlying target structures are presented. Being didactic tools mediating the mind that can be assigned psychological status, concepts are expected to be materialized to learners by means of a Schema of a Complete Orienting Basis of an Action (SCOBA), which provides learners with a holistic cognitive map mediating their minds. Teachers assist learners in studying the SCOBA and comprehending concepts involved (Gal’perin, 1989, 1992; Haenen, 1996; Lantolf & Poehner, 2014; Negueruela, 2008). Having comprehended semantic and pragmatic concepts underlying target structures materialized in the SCOBA, students proceed to two subsequent stages warranting verbalization of concepts: stages of overt speech and
covert speech. Provided with primary language data with instances of target language structures, learners are required to explicate their understanding of target forms in those samples in relation to learnt concepts to their peers and themselves at stages of overt speech and covert speech respectively (Gal’perin, 1969; Haenen, 1996; Lantolf & Poehner, 2014). Having gained mastery of concepts underlying target language structures at the two stages of verbal action, students eventually reach the final stage of STI: the mental stage, where overt and covert speeches have been transformed into inner speech.

**Cognitive Grammar**

Cognitive grammar (CG) provides the linguistic framework for pedagogical grammar method under investigation in the study. The English tense system is conceived by cognitive grammar using an epistemic model (see Figure 1 below). The two English tenses are exploited to delineate reality (R), which is metaphorically compared to growing cylinder with growth occurring continuously at its leading face representing current reality and human beings as conceptualizers (C) (Langacker, 2008). The portion of current reality accepted by conceptualizers to be real is known as immediate reality (IR) whilst the remaining portion of reality accepted by conceptualizers to be real is known as conceived reality (Rc) or non-immediate reality (ibid.). At any moment, a conceptualizer directs his/her attention to one particular situation, be it a state or an event; that particular situation is known as profiled occurrence (PO) (Langacker, 2011). Profiled occurrence necessarily exists within the conceptualizer’s immediate scope (IS), which is metaphorically conceived as an onstage region of the conceptualizer’s attention, within the maximal scope (MS), which comprises the overall content (ibid.).

![Figure 1. An Epistemic Model.](image)


The present and past tenses delineate profiled occurrences within immediate reality and non-immediate reality respectively (Langacker, 2011). Notwithstanding not delineating profiled occurrences within immediate reality and non-immediate reality temporally respectively, peripheral usages of the present tense, videlicet timeless truths and scheduled future, and those of
the past tense, videlicet imaginative conditionals and social distancing, are applied when profiled occurrences enter conceptualizers’ immediate reality and non-immediate reality virtually respectively (ibid.). The entirety of the English tense system can thereby be explicated using the semantic concept of epistemic reality, under which all usages of the present and past tenses are subsumed.

Application of STI and CG

STI has been implemented in L2 classrooms with various target structures of distinct target languages all over the globe. The most prominent study of implementation of STI in L2 classrooms was conducted by Negueruela (2003) in a university-level Spanish-as-a-foreign-language course with Spanish tense, aspect, and modality as target structures. Having been presented an accurate and complete picture of scientific concepts underlying target structures through flow charts, students, who had already been exposed to target structures prior to the course, engaged in verbalization tasks requiring them to make use of learnt concepts to explicate language use in given primary language data (ibid.). Participants were discovered from verbalization and performance data to reveal more accurate conceptual understanding as well as more accurate written and oral production of target structures after instruction (ibid.). More importantly, despite having experienced conflicts between such instruction and prior instruction, participants recognized the value of distinct instructional procedures and were capable of developing their learner agency on the whole (ibid.). Positive results of Negueruela’s study motivated subsequent empirical studies on STI.

Akin to Negueruela’s study, all antecedent empirical studies exploring implementation of STI involved pedagogical procedures of materialization, verbalization, and internalization. van Compernolle (2012), Kim (2013), and Zhang (2014) utilized STI in instruction on the sociopragmatics of L2 French, L2 sarcasm, and topicalization in Chinese in laboratory settings respectively whereas Yanez-Prieto (2008) and Lai (2012) applied STI to instruction on Spanish verbal aspects and the Chinese temporal system in intact classrooms respectively. STI has been found to manage to assist foreign language learners in overcoming pre-understanding on the basis of rules of thumb and enhancing their conceptual development as well as accuracy of production of target structures. All the same, some teachers and students appear to possess reservations about STI out of unfamiliarity, and more studies ought to be conducted in authentic classrooms in lieu of laboratory settings.

Attempts have also been made to capitalize upon STI in instruction on the English tense and aspect system, which is closely associated with the target structure of the present study, in tertiary L2 classrooms (e.g. Ganem-Gutierrez & Harun, 2011; Harun, Abdullah, Ab Wahab, & Zainuddin, 2017, 2019; Harun, Massari, & Behak, 2014; Ng & Zhao, 2017; Pohner & Infante, 2015, 2017). Having analyzed learners’ individual and dyadic verbalization during and after instruction, Ganem-Gutierrez and Harun (2011), Harun et al. (2014), as well as Ng and Zhao (2017) have discovered that learners’ usage of metalanguage in the course of verbalization serves as a regulatory tool revolutionizing or deepening their conceptual understanding of the language system. Recognizing the value of mediator guidance, Pohner and Infante (2015, 2017) have found that expert scaffolding is required for learners’ comprehension of concepts presented. In spite of efficacy of STI in enhancing students’ grammatical production of the English tense and aspect system demonstrated in the aforementioned studies, more studies ought to be conducted in non-tertiary L2 classrooms in a bid to probe into applicability of STI in disparate contexts. A major difference between tertiary and non-tertiary settings is that General English is mainly taught at a
non-tertiary level while English for Specific Purposes constitutes the focus of English language learning at a tertiary level.

In contrast, only have limited attempts been made to employ CG in L2 grammar instruction; the overwhelming majority of those studies targeted on structures posing considerable challenge to learners. For instance, Tyler, Mueller, and Ho (2010a, 2010b) conducted two experimental studies of implementation of CG in instruction on English modals and prepositions in L2 classrooms respectively by comparing impacts of instruction involving CG-motivated instructional materials and that predominantly involving explication of grammar using rules of thumb on participants’ learning outcomes. Providing learners with precise definitions of target structures as well as systematic accounts of interrelations amongst related language items, CG-based instruction is discovered to effectuate learners’ more significant gains in accuracy of production as well as more profound improvement in understanding of target structures (Tyler et al., 2010a, 2010b). Superiority of CG-motivated instruction over traditional instruction in terms of pedagogical efficacy has thereby been substantiated.

CG has also been adopted in instruction on the English tense and aspect system in L2 classrooms, yet results appear much less encouraging than the aforementioned ones. Bielan and Pawlak (2013) as well as Kermer (2016) implemented CG-inspired instruction to teach English tenses and aspects in L2 classrooms and compared its impacts on learners’ receptive and productive knowledge on target structures with those of instruction based on descriptive grammar rules. Unlike Tyler, Mueller, and Ho’s studies, neither was CG-informed instruction found to be significantly more profitable for nor efficacious in enhancing learners’ receptive or productive knowledge on target structures in the two studies concerning tense and aspect (Bielan & Pawlak, 2013; Kermer, 2016). One plausible explication for low pedagogical efficacy of CG-informed instruction in those two studies was that by no means were those CG instructional materials congruent with principles of CG even though they were claimed to be informed by CG. More specifically, barely was each of the two tenses treated as one symbolic unit, but distinct usages of the present and past tenses were presented as independent rules as in descriptive grammar in CG treatments of both studies (Bielan & Pawlak, 2013; Kermer, 2016). For this reason, more studies of application of CG in L2 grammar instruction, in particular instruction on English tense and aspect, using instructional materials truly reflecting CG principles are doubtlessly necessary to evaluate pedagogical efficacy of CG-based instruction accurately.

**Methodology**

The present paper is a segment of a larger study aiming at exploring application of a pedagogical method integrating STI with CG to instruction on the English tense system in an L2 classroom in a Hong Kong secondary school. A mixed-method approach was employed in analysis of impacts of instruction integrating STI with CG on students’ grammatical production of the English present and past tenses. Students’ grammatical production is the dependent variable whereas participants’ level of English proficiency and concepts were the between-participant and within-participant independent variable respectively.

**Participants**

Being a classroom-based study addressing a problem with grammar pedagogy confronted by L2 educators, the research was conducted in an L2 classroom in a Hong Kong secondary school. At a junior secondary level, participants were presumed to have been taught the English tense system and basically mastered the overwhelming majority of usages of the system, so the goal of
the instruction was to build upon their prior knowledge, advance their conceptual development, and eventually promote their linguistic development. Twenty-nine secondary, three students aged between 13 and 14, with 10 male and 19 female, were recruited for the study.

A total of 13 students were selected by means of critical case sampling and snowball sampling for exit focus group interviews, each of which comprised three to five participants. A number of students with varying levels of English proficiency were first selected by the researcher, and they were subsequently encouraged to invite their peers to join those interviews (Noy, 2008; Teddlie & Tashakkori, 2009; Teddlie & Yu, 2007; Watts & Ebbutt, 1987). In this vein, opinions of students with distinct achievement levels could be assembled whereas more responses could be generated from closeness amongst interviewees.

**Instructional Instruments**

Instrumental instruments adopted in instructional sessions comprised a collection of SCOBAs and verbalization tasks. SCOBAs of the study were designed on the basis of Langacker’s (2008) epistemic model and adapted from materials developed by Langacker (2011). They comprised a simplified epistemic model (see Figure 2 below), which introduces the concept of epistemic reality, three didactic charts presenting the four semantic concepts underlying the English tense system (see Figures 3a-3c below), videlicet temporal immediacy, virtual immediacy, temporal non- immediacy, and virtual non-immediacy, as well as mapping various usages of the two tenses onto the four concepts, and four diagrams elucidating the four semantic concepts at length (see Figures 4a-4d below). In-class verbalization tasks and after-class written verbalization tasks required participants to explicate their understanding of tenses of verbs in primary language data in relation to the four learnt concepts. Tables 1 and 2 below present an instance of an in-class verbalization task and an after-class written verbalization task respectively.

![Figure 2. A Simplified Epistemic Model.](image)
Figure 3. Three didactic charts.
Table 1. A Sample In-Class Verbalization Task

Explain the tense of the underlined verb in each sentence to your partner using the given charts and diagrams.

There is a large house on the corner.


Table 2. A Sample After-Class Written Verbalization Task

Explain the tense of the underlined verb in each sentence. You may use diagrams if necessary.

The car belongs to Bill.


Data Collection Instruments

Data were collected through a pre-test and a post-test, each of which comprised two parts, videlicet controlled grammatical production and free grammatical production, respectively. For all elicitation of participants’ grammatical production, tests of controlled production and free production vary in terms of length of input provided and length of expected response (Purpura, 2004). Controlled production tests generally provide participants with more input and necessitate shorter responses whereas free production tests generally provide participants with less input and necessitate longer responses (ibid.). Measuring participants’ explicit and implicit grammatical knowledge respectively, both controlled and free grammatical production were incorporated into the tests (ibid.).

The controlled grammatical production test comprised a total of 24 sentential level gap-filling test items. Presenting input in the form of a sentence as well as the base form of a verb as a prompt.
and requiring learners to fill gaps, the test was utilized to gauge learners’ controlled production of English present and past tenses in the study (Hughes, 2003; Purpura, 2004). Table 3 shows a sample sentential level gap-filling test item. In the sample item, the belief is a present state, so the present tense of the verb ‘believe’ should be used to fill the gap.

Table 3. A Sample Sentential Level Gap-Filling Test Item

Fill in the blanks with correct forms of verbs in brackets.
Example:
We ____________________ (believe) this point is vital for supporting adolescent struggling readers.
Source of sentence: Davies (2008)

There being four target semantic concepts, the test comprised an equal token of four test items targeting at each concept and eight filler items, which manifested use of verb forms other than the present simple and the past simple, for fair comparisons of students’ grammatical performance in relation to the four concepts. This made up a total of 24 test items. Lexical aspects of verbs being one controlled variable of the study, all verbs present in primary language data in the test were state or achievement verbs to ensure that students’ grammatical performance was independent of lexical aspects of verbs (Salaberry & Shirai, 2002; Vendler, 1967; Yule, 1998).

The free grammatical production test was a written production task. Presenting learners with a prompt eliciting distinct usages of the present and past tenses and requiring them to write an essay, which simulates authentic language use, the test was an avenue for assessing learners’ free production of English present and past tenses in the study (Purpura, 2004). Table 4 shows a sample written production item. In the sample item, the first two bullet points of the prompt elicit production of the present tense whilst the last two bullet points elicit production of the past tense.

Table 4. A Sample Written Production Item

You are a secondary one student who has attended school for two weeks, and you are asked to write a short paragraph of about 100 words describing your secondary school and secondary school life.
Please do the following things in the paragraph:
• Briefly describe your secondary school (e.g. the name of the school, school facilities, etc.)
• Write about your everyday life and habits as a secondary school student
• Compare your secondary school life with your primary school life in the past
• Imagine how your secondary school life might be different if you went to a different secondary school

Written production tasks were assessed on the basis of suppliance in obligatory contexts (SOC) (e.g. Bardovi-Harlig, 2000; Brown, 1973); a trained second rater was invited to assess 30% of selected scripts, and the simple percentage agreement between the researcher and the second rater was 92% (Mackey & Gass, 2016).

Three sets of tests had been prepared in advance and were randomly assigned to participants as pre-test and post-test (Mackey & Gass, 2016); such counterbalancing enhanced internal validity of the research. Participants’ test scores were analyzed both descriptively and inferentially to compare students’ performance in the two tests.

Focus group interviews garnered elaborate responses from participants regarding their perceptions of the entirety of the learning experience, which illuminated impacts of the pedagogical approach on their grammatical production. An interview protocol with 11 questions
having been predetermined on the basis of Lee’s (2012) study, standardized open-ended interviews, which are opined to be semi-structured, were applied to the study to enhance comparability of responses and facilitate data analysis (Patton, 1980). Notwithstanding existence of a list of questions, the researcher remained flexible and asked follow-up questions when necessary during interviews. Interviews were conducted in Cantonese, students’ first language, so that detrimental impacts of learners’ level of English proficiency on quality and quantity of data provided could be eliminated (Mackey & Gass, 2016). Interview data were analyzed via qualitative content analysis, where interviews were transcribed, translated, pre-coded, and coded (Dornyei, 2007). Content of interviews was interpreted and reported by means of identification of recurring themes and extraction of quotes providing evidence for each theme (Krippendorp, 2004).

**Procedures**

The entirety of the study lasted for seven weeks. Participants signed a consent form and completed an entry questionnaire adapted from Li, Zhang, Tsai, and Puls (2014), which assembled information on their language proficiency, language history, and experience of learning English tenses, as well as a pre-test in the first week.

Instructional sessions were conducted by the researcher on a weekly basis from the second week to the fifth week with each session focusing on one of the four semantic concepts underlying the English tense system. Each instructional session lasted for 25 minutes and followed pedagogical procedures of STI put forward by Gal’perin (1969) and outlined by Haenen (1996). The target semantic concept of each instructional session was presented and elucidated to students with assistance of SCOBAs and primary language data at the orienting and material(ized) stages. After the two stages of materialization of concepts, students were required to explicate tenses of verbs in primary language data in relation to the learnt concept to their peers and themselves at stages of overt speech and covert speech respectively; written verbalization tasks were also provided for completion after class. Both in-class verbalization tasks and after-class written verbalization tasks paved way for internalization of concepts at the mental stage. An equal number of sentences was presented to students in each instructional session to ensure that the same amount of input of the four concepts was provided for students for a fair comparison.

A post-test was administered momentarily after the final instructional session in the fifth week to look into impacts of the instruction on participants’ grammatical production of the English tense system. Focus group interviews were conducted in the last two weeks of the research period.

**Results and Discussion**

Presentation of findings is divided into three levels: overall analysis, between–participant analysis, where level of English proficiency is an independent variable, and within–participant analysis, where the target concepts under investigation constitute an independent variable. Participants were divided into three groups of high, mid, and low levels of English proficiency on the basis of their scores of daily writing assignments for between–participant analysis.

**Controlled Grammatical Production**

**Overall Analysis**

A paired-samples t-test was conducted to examine impacts of the intervention on participants’ controlled grammatical production through comparison between participants’ controlled production in sentential level gap-filling tests in the pre-test and the post-test. Descriptive statistics
of participants’ overall controlled production are presented numerically and graphically in Table 5 and Figure 5 respectively.

Table 5. Descriptive Statistics of Participants’ Overall Controlled Grammatical Production

<table>
<thead>
<tr>
<th>Test</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>BCa bootstrapped 95% CI of mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>29</td>
<td>5</td>
<td>16</td>
<td>9.45</td>
<td>2.41</td>
<td>.36</td>
<td>.73</td>
<td>[8.55, 10.31]</td>
</tr>
<tr>
<td>Post-test</td>
<td>29</td>
<td>6</td>
<td>15</td>
<td>10.14</td>
<td>2.37</td>
<td>.34</td>
<td>-.42</td>
<td>[9.24, 10.93]</td>
</tr>
</tbody>
</table>

Figure 5. Descriptive Statistics of Participants’ Overall Controlled Grammatical Production.

Results of the t-test ($t_{28} = 1.24$, $p = .22$, 95% BCa CI = [-.45, 1.72], $d = 12.20$) manifest no statistical difference in participants’ controlled grammatical production between the pre-test and the post-test in that the $p$-value was statistically insignificant at a .05 level albeit a large effect size of 12.20. This implies that participants’ overall controlled production in the pre-test and the post-test differed by 12.20 SDs. Such findings possess a propensity to suggest that the instruction potentially possesses a considerable magnitude of effect on students’ overall controlled production of the English tense system, yet hardly can such an effect be reflected in results of the t-test owing to a small sample size of the study (Larson-Hall, 2016). These are consistent with those of antecedent studies on STI (e.g. Negueruela, 2003; Ng & Zhao, 2017) as well as integration of STI and CG (e.g. Lee, 2012; White, 2012), which discovered significant impacts of such intervention on learners’ grammatical production.
**Between-participant Analysis**

Three paired-samples t-tests were conducted to compare controlled grammatical production in sentential level gap-filling tests in the pre-test and the post-test for participants in the three proficiency groups. Descriptive statistics of controlled production of participants in distinct proficiency groups are presented numerically and graphically in Table 6 and Figure 6 respectively, and results of t-tests are displayed in Table 7.

**Table 6. Descriptive Statistics of Controlled Grammatical Production of Participants at Distinct Levels of English Proficiency**

<table>
<thead>
<tr>
<th>Level of English proficiency</th>
<th>N</th>
<th>Mean (SD) of pre-test</th>
<th>Mean (SD) of post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>10</td>
<td>10.20 (3.16)</td>
<td>10.90 (2.73)</td>
</tr>
<tr>
<td>Mid</td>
<td>9</td>
<td>8.78 (2.22)</td>
<td>9.33 (1.66)</td>
</tr>
<tr>
<td>Low</td>
<td>10</td>
<td>9.30 (1.64)</td>
<td>10.10 (2.51)</td>
</tr>
</tbody>
</table>

**Figure 6.** Descriptive Statistics of Controlled Grammatical Production of Participants at Distinct Levels of English Proficiency.
Table 7. Paired-Samples T-Tests of Controlled Grammatical Production of Participants at Distinct Levels of English Proficiency

<table>
<thead>
<tr>
<th>Level of English Language Proficiency</th>
<th>Mean Difference</th>
<th>df</th>
<th>t</th>
<th>p</th>
<th>95% BCa CI</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>.70</td>
<td>9</td>
<td>.94</td>
<td>.37</td>
<td>[-.83, 2.11]</td>
<td>1.15**</td>
</tr>
<tr>
<td>Mid</td>
<td>.56</td>
<td>8</td>
<td>.49</td>
<td>.64</td>
<td>[-1.58, 2.71]</td>
<td>.69</td>
</tr>
<tr>
<td>Low</td>
<td>.80</td>
<td>9</td>
<td>.74</td>
<td>.48</td>
<td>[-1.20, 2.82]</td>
<td>.65</td>
</tr>
</tbody>
</table>

Note. *p < .05. **d > .8

Results of t-tests indicate no statistical difference in controlled grammatical production between the pre-test and the post-test for participants at all three levels of English proficiency in that all p-values were statistically insignificant at a .05 level albeit a large effect size for the high proficiency group (Cohen’s d = 1.15). This implies that controlled production of this group of participants in the pre-test and the post-test differed by 1.15 SDs. It thereby appears that students at a higher level of English proficiency ameliorated more profoundly in their grammatical comprehension of the English tense system after the instruction than their counterparts did; one plausible reason is more able students’ better mastery of their L2 as a mediational tool. Language is an effectual tool for metacognitive mediation as suggested in Swain’s (2006) concept of languaging and evidenced in empirical studies (e.g. Ganem-Gutierrez & Harun, 2011; Harun et al., 2014). Some participants in the high proficiency group were thereby capable of learning those concepts simply by looking at words without resorting to diagrams as expressed in focus group interviews, which meant that they better mastered English as a tool to mediate their minds. In contrast, those in the low proficiency demanded more time for and possessed more difficulty in comprehension of those concepts plausibly by reason of their low level of English proficiency; better comprehension of concepts during materialization enabled participants at a high level of English proficiency to better internalize those concepts and apply them to grammatical production.

“… [these diagrams] are sometimes unnecessary.”

(Focus Group 1, Student 3, High-Level)

“It takes more time for comprehension.”

(Focus Group 1, Student 1, Low-Level)

Within-participant Analysis

Four paired-samples t-tests were conducted to compare participants’ controlled grammatical production of the four concepts in sentential level gap-filling tests in the pre-test and the post-test. Descriptive statistics of participants’ controlled production of distinct concepts are presented numerically and graphically in Table 8 and Figure 7 respectively, and results of t-tests are displayed in Table 9.
Table 8. Descriptive Statistics of Participants’ Controlled Grammatical Production of Distinct Concepts

<table>
<thead>
<tr>
<th>Concept</th>
<th>N</th>
<th>Mean (SD) of Pre-test</th>
<th>Mean (SD) of Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal immediacy</td>
<td>29</td>
<td>2.24 (.87)</td>
<td>2.48 (.95)</td>
</tr>
<tr>
<td>Virtual immediacy</td>
<td>29</td>
<td>2.90 (.94)</td>
<td>2.90 (1.08)</td>
</tr>
<tr>
<td>Temporal non-immediacy</td>
<td>29</td>
<td>2.38 (.98)</td>
<td>2.52 (.95)</td>
</tr>
<tr>
<td>Virtual non-immediacy</td>
<td>29</td>
<td>1.93 (1.39)</td>
<td>2.24 (1.12)</td>
</tr>
</tbody>
</table>

Figure 7. Descriptive Statistics of Participants’ Controlled Grammatical Production of Distinct Concepts.

Table 9. Paired-Samples T-Tests of Participants’ Controlled Grammatical Production of Distinct Concepts

<table>
<thead>
<tr>
<th>Concept</th>
<th>Mean Difference</th>
<th>df</th>
<th>t</th>
<th>p</th>
<th>95% BCa CI</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal immediacy</td>
<td>.24</td>
<td>28</td>
<td>1.19</td>
<td>.24</td>
<td>[-.14, .66]</td>
<td>2.12**</td>
</tr>
<tr>
<td>Virtual immediacy</td>
<td>.00</td>
<td>28</td>
<td>.00</td>
<td>1.00</td>
<td>[-.55, .48]</td>
<td>.00</td>
</tr>
<tr>
<td>Temporal non-immediacy</td>
<td>.14</td>
<td>28</td>
<td>.54</td>
<td>.60</td>
<td>[-.38, .59]</td>
<td>3.39**</td>
</tr>
<tr>
<td>Virtual non-immediacy</td>
<td>.31</td>
<td>28</td>
<td>1.01</td>
<td>.32</td>
<td>[-.31, .86]</td>
<td>.81**</td>
</tr>
</tbody>
</table>

Note. *p < .05. **d > .8
Results of t-tests signify that the instruction exerted much impact on participants’ controlled grammatical production of all concepts save virtual immediacy albeit a lack of statistical difference observed in mean scores of any of the four concepts. The effect sizes (Cohen’s $d$) of the three concepts were large, yet $p$-values of the concepts were statistically insignificant at a .05 level. This could be accounted for by both instructional order and relative abstractness of the concept. Concept taught in latter instructional sessions, when learners had already been more familiar with pedagogical activities in instructional sessions, and less abstract ones, videlicet temporal in lieu of virtual concepts, were generally more easily comprehended by learners (Langacker, 2011). Being an abstract concept taught in the second instructional session, the concept of virtual immediacy might be incomprehensible to participants at that stage and even confused their understanding, so no statistically significant improvement in controlled production of such a concept was observed in the post-test. A participant of a focus group interview commented that only when she had reached later instructional sessions did her mind begin to become clearer about concepts taught, so more significant improvement was observed in participants’ grammatical production of concepts taught in later instructional sessions.

“(Interviewer: After which lesson did your mind become much clearer?) The second and third lessons.”

(Focus Group 3, Student 2, Mid-Level)

### Free Grammatical Production

#### Overall Analysis

A paired-samples t-test was conducted to explore impacts of the intervention on participants’ free grammatical production via comparison between participants’ free production in written production tasks in the pre-test and the post-test. Descriptive statistics of participants’ overall free production are presented numerically and graphically in Table 10 and Figure 8 respectively.

<table>
<thead>
<tr>
<th>Test</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Bca bootstrapped 95% CI of mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>29</td>
<td>29</td>
<td>100</td>
<td>77.63</td>
<td>18.47</td>
<td>-1.12</td>
<td>.78</td>
<td>[70.65, 83.68]</td>
</tr>
<tr>
<td>Post-test</td>
<td>29</td>
<td>46</td>
<td>100</td>
<td>81.55</td>
<td>15.90</td>
<td>-.69</td>
<td>-.37</td>
<td>[75.81, 87.51]</td>
</tr>
</tbody>
</table>
Results of the t-test ($t_{28} = 1.04$, $p = .31$, 95% BCa CI = [-4.88, 11.86], $d = 1.08$) exhibit no statistical difference in participants’ free grammatical production between the pre-test and the post-test. The $p$-value was statistically insignificant at a .05 level albeit a large effect size of 1.08. This implies that participants’ overall free production in the pre-test and the post-test differed by 1.08 SDs. Akin to findings of participants’ overall controlled production, those of participants’ overall free production are apt to suggest that the instruction potentially possesses a considerable magnitude of effect on students’ grammatical production of the English tense system, yet the effect was plausibly limited by the small sample size.

**Between-participant Analysis**

Three paired-samples t-tests were conducted to compare free grammatical production in written production tasks in the pre-test and the post-test for participants in the three proficiency groups. Descriptive statistics of free production of participants in distinct proficiency groups are presented numerically and graphically in Table 11 and Figure 9 respectively, and results of t-tests are displayed in Table 12.

**Table 11. Descriptive Statistics of Free Grammatical Production of Participants at Distinct Levels of English Proficiency**

<table>
<thead>
<tr>
<th>Level of English proficiency</th>
<th>N</th>
<th>Mean (SD) of pre-test</th>
<th>Mean (SD) of post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>10</td>
<td>83.02 (12.09)</td>
<td>87.83 (15.27)</td>
</tr>
<tr>
<td>Mid</td>
<td>9</td>
<td>82.33 (13.70)</td>
<td>87.31 (12.27)</td>
</tr>
<tr>
<td>Low</td>
<td>10</td>
<td>68.00 (24.31)</td>
<td>70.09 (14.00)</td>
</tr>
</tbody>
</table>
Results of t-tests reveal no statistical difference in free grammatical production between the pre-test and the post-test for participants at all three levels of English proficiency in that all p-values were statistically insignificant at a .05 level albeit large effect sizes for the high and mid proficiency groups. It thereby appears that students at a higher level of English proficiency ameliorated more profoundly in their grammatical comprehension of the English tense system after the instruction than their counterparts did. Such findings are in sync with those of participants’ controlled production and could similarly be explicated by more able students’ better use of the English language as a mediational tool for materialization and internalization of concepts, which was presumed to be profitable for both controlled and free production.
**Within-participant Analysis**

Two paired-samples t-tests were conducted to compare participants’ free grammatical production of the two tenses in written production tasks in the pre-test and the post-test. Descriptive statistics of participants’ free production of distinct tenses are presented numerically and graphically in Table 13 and Figure 10 respectively, and results of t-tests are displayed in Table 14.

### Table 13. Descriptive Statistics of Participants’ Free Grammatical Production of Distinct Tenses

<table>
<thead>
<tr>
<th>Tense</th>
<th>N</th>
<th>Mean (SD) of Pre-test</th>
<th>Mean (SD) of Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present tense</td>
<td>29</td>
<td>88.14 (15.78)</td>
<td>91.79 (14.48)</td>
</tr>
<tr>
<td>Past tense</td>
<td>29</td>
<td>62.06 (33.00)</td>
<td>65.46 (29.81)</td>
</tr>
</tbody>
</table>

*Figure 10. Descriptive Statistics of Participants’ Free Grammatical Production of Distinct Tenses.*

### Table 14. Paired-Samples T-Tests of Participants’ Free Grammatical Production of Distinct Tenses

<table>
<thead>
<tr>
<th>Concept</th>
<th>Mean difference</th>
<th>df</th>
<th>t</th>
<th>p</th>
<th>95% BCa CI</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present tense</td>
<td>3.65</td>
<td>28</td>
<td>1.05</td>
<td>.30</td>
<td>[-2.56, 11.88]</td>
<td>1.99**</td>
</tr>
<tr>
<td>Past tense</td>
<td>4.36</td>
<td>28</td>
<td>.56</td>
<td>.58</td>
<td>[-10.75, 19.30]</td>
<td>.75</td>
</tr>
</tbody>
</table>

*Note. *p < .05. **d > .8*
Results of t-tests demonstrate that the instruction exerted much impact on participants’ free grammatical production of the present tense albeit a lack of statistical difference observed in mean scores of both tenses in that the effect size (Cohen’s $d$) of the present tense was frightfully large, yet $p$-values of both tenses were statistically insignificant at a .05 level. In addition, the accuracy rate of the present tense was higher than that of the past tense in both the pre-test and the post-test. Both a large effect size of free production of merely the present tense and higher accuracy rates of the present tense in both tests could be accounted for by from a morphological perspective. The present tense and past tense of English verbs being inflected only for third-person singular subjects and obligatorily carrying inflections respectively, the English present tense and past tense are said to possess unmarked and marked morphology respectively (Huddleston, 2002). For such a reason, the past tense is construed as morphologically more challenging to L2 learners, so it is for the accuracy rate of the past tense to be lower than that of the present tense in free production. Even if learners are aware of a need to delineate situations in the past tense, a failure in provision of the past tense in actual language production, or morphological variability, may be observed (Haznedar & Schwartz, 1997; Ionin & Wexler, 2002; Lardiere, 1998a, 1998b; Law, 2005; McArthur, 2002).

Integration of STI and CG targeting at the semantics in lieu of morphology of the English tense system, never could learners’ morphological difficulty be directed addressed.

**Limitations**

Two predominant limitations of the present study are the small sample size and limited instructional time.

First and foremost, solely one class of students recruited for the current study, barely was the sample size of the study large enough for effect of the instruction to be reflected in inferential statistics and for findings to be generalized even to classrooms in the same educational contexts, let alone those in other contexts. It is thereby suggested to conduct future studies in multiple classrooms in distinct schools with the hope of yielding more informative findings of impacts of instruction integrating STI with CG on grammatical performance of students at distinct contexts. Studies may even compare impacts of the pedagogy in various contexts, such as Asian and western contexts, to observe whether pedagogical efficacy of the pedagogy is influenced by contextual factors. Another pitfall arising from the small sample size is that disparities in level of English proficiency amongst students were insignificant, so was the value of the between-participant analysis in the study. Should the study be conducted in multiple contexts, it will also be more likely for students at more diverse levels of English proficiency to be recruited. When categorizing students into proficiency groups, it is more desirable to capitalize upon General English proficiency tests in lieu of merely scores of daily writing assignments to categorize students for more accurate measurement of students’ level of English proficiency.

Another limitation of the study is limited time devoted to instruction on the four target concepts. Sufficient opportunities for mediation are integral for internalization of concepts, yet participants in the current study experienced inadequate mediation owing to limited instructional time. More importantly, segmentation of instruction on the entirety of the English tense system into four instructional sessions, which is an undesirable pedagogical practice in STI, also resulted from limited instructional time. It is thereby recommended to cover the entirety of the English tense system in one instructional session and lengthen the total instructional time so that students can be provided with a coherent learning experience with more opportunities for mediation.
Conclusion

By and large, attempting to expand the body of literature on application of STI and CG to authentic L2 classrooms and illuminate future development of L2 grammar pedagogy of English tenses, the present paper aims at investigating impacts of instruction integrating STI with CG on students’ grammatical production. On the whole, barely does integration of STI and CG ameliorate students’ grammatical production statistically, yet the effect size of the pedagogy on students’ grammatical production was large. Such findings suggest that the instruction potentially possesses a considerable magnitude of effect on students’ overall grammatical production of the English tense system, yet hardly can such an effect be reflected in results of the t-test owing to a small sample size of the study.

Offering a pedagogical alternative for grammar instruction, integration of STI and CG complements the inadequacy of traditional pedagogical approaches, yet its pedagogical efficacy in ameliorating students’ grammatical production may be limited by factors such as the length of instructional time and instructional order. This provides implications for both educators and researchers.

From educators’ perspective, the two aforementioned factors that may limit pedagogical efficacy of the pedagogy ought to be taken into consideration in lesson design so that the pedagogical efficacy of the pedagogical approach can be maximized for learners’ sake. For instance, limited instructional time detracting from pedagogical efficacy of the pedagogy as a result of insufficient opportunities for mediation, capability to devote sufficient time to instruction of one target structure is a premise for successful implementation of the pedagogy in any L2 classrooms. Should such pedagogy be put into practice in L2 classrooms, it is also suggested to plan the instruction on all concepts in the same instructional sessions in lieu of distinct sessions so that it will be easier for students to comprehend the interconnection amongst concepts and acquire comparable amounts of knowledge on distinct concepts.

From researchers’ perspective, several future research ideas are put forward for the sake of following up the present study and looking into implementation of instruction integrating STI with CG more deeply. More specifically, in light of identified limitations of the current study, it is recommended to conduct a replication study with ameliorated research design and pedagogical design, such as a larger sample and lengthening of instructional time for provision of time for learners to internalize learnt concepts. Application of the pedagogy can also be extended to instruction on other target structures, such as interactions between tense and aspect, to identify target structures better fitting such pedagogy. It is hoped that more studies integrating STI as a pedagogical framework with CG as a linguistic framework can be conducted in the future so that possibilities of integration of the two theories can be better illuminated, and such an integration can also be promoted in the field of L2 grammar instruction.
References


