



Contents lists available at ScienceDirect

Journal of Second Language Writing

journal homepage: www.elsevier.com/locate/jslw

Challenging the connection between task perceptions and language use in L2 writing: Genre, cognitive task complexity, and linguistic complexity

Hyung-Jo Yoon

California State University, Northridge, United States

ARTICLE INFO

Keywords:

Cognitive task complexity
Genre
Second language writing
Linguistic complexity
Written production
Syntactic complexity

1. Introduction

There has been much research on the effects of task condition on second language (L2) writers' language use (e.g., Frear & Bitchener, 2015; Johnson et al., 2012; Kormos, 2011; Lee, 2020; Ong, 2014; Révész et al., 2017). These studies have generally explored if task manipulations elicit different linguistic features from L2 learners, as predicted by theoretical frameworks such as the cognition hypothesis (Robinson, 2001, 2005) and the limited attentional capacity model (Skehan, 1998). They were originally developed to explain the effects of task manipulations on speech production, so the focus of L2 writing research has been on the transfer of their predictions to written production (Byrnes & Manchón, 2014). However, the findings of previous writing studies have not offered clear evidence of the predicted effects of task manipulations on language in written discourse (see Johnson, 2017, for a review).

In this regard, it has been noted that the two modalities (oral and written) differ fundamentally and that it may not be valid to apply the cognitive complexity hypotheses to writing research directly (Frear & Bitchener, 2015; Johnson et al., 2012). Specifically, the underlying assumption of cognitive task complexity involves the allocation of limited attentional resources during spoken language production, whereas, as a recursive process, writing would be less constrained by such cognitive limitations (Hayes, 1996, 2012). Yuan and Ellis (2003) also noted that writing activities tend to be less pressured than speaking due to the extended availability of planning and monitoring throughout the process. However, Vasylets et al. (2017) found that the very nature of reduced constraints in written

E-mail address: hyungjo.yoon@csun.edu.

<https://doi.org/10.1016/j.jslw.2021.100857>

Received 6 December 2020; Received in revised form 6 November 2021; Accepted 8 November 2021

Available online 29 November 2021

1060-3743/© 2021 Published by Elsevier Inc.

performance resulted in greater task complexity effects on language features in writing than in speaking (i.e., increased linguistic complexity for a more complex writing task in accordance with the cognition hypothesis). This study is grounded in the cognitive model of writing proposed by Hayes (1996). While originally developed to explain first language (L1) writing, his model offers a detailed description of recursive writing processes that L2 writers would also experience. It includes the two major factors of the task environment and the individual. The latter is composed of four components that interact with each other: motivation/affect, cognitive processes, working memory, and long-term memory. The long-term memory component, including task schemas and genre knowledge, is particularly relevant to the focus of this study.

Task-based writing studies have examined simpler and more complex versions of the same task to inform task sequencing decisions, although genre has also been explored quite extensively as a writing task feature (e.g., Qin & Uccelli, 2016; Ruiz-Funes, 2015; Yang, 2014). Findings of these studies have shown that genre is actually one of the few variables that induces clear differences in language use. Specifically, L2 writers tend to use more syntactically complex structures for argumentative than descriptive or narrative tasks, and these results have often been interpreted as evidence of greater cognitive demands imposed by argumentative writing than other genres (discussed in more detail below). However, an underlying motivation for varying syntactic complexity may not necessarily be the cognitive complexity of a task, and as Housen et al. (2019) argued, clear distinctions should be made between different notions of complexity. Specifically, cognitive complexity and difficulty are regarded as relative complexity, while linguistic complexity is regarded as absolute (structural) complexity. Although a considerable amount of research has attended to linguistic complexity of learner writing across genres, little research has been conducted on how learners perceive writing genres.

Given these gaps, there is a need to clarify the cognitive demands of different genres and the meaning of genre effects on language. To achieve these aims, this study examines L2 learners' production and perceptions of four tasks. The writing tasks targeted in the study involve timed narrative and argumentative genres (telling an interesting story to entertain readers and making arguments to convince readers) within which a level of task complexity is manipulated in terms of the availability of supporting ideas. In this paper, *task type* is used as a broader term for writing tasks manipulated by genre or task complexity. The learners' task perceptions are collected immediately after their writing performance via a self-rating questionnaire. Their task perceptions and writing performance are examined together to explain the motivation for linguistic changes in writing.

2. Literature review

2.1. Cognitive task complexity in L2 writing

It has been argued that task complexity, defined as the "attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner" (Robinson, 2001, p. 29), influences the amount of attentional and cognitive resources available for language structures during task performance. Thus, manipulations of task complexity are expected to generate different levels of cognitive demands in the conceptualization stage, which in turn lead to changes in the complexity of linguistic forms (Robinson, 2005; Skehan, 1998). Several task variables are hypothesized to affect the allocation of attentional resources for linguistic complexity (e.g., here-and-now, number of elements, and conceptual demands). In the context of L2 writing, much attention has been given to the effect of conceptual demands on language use (e.g., Kormos, 2011; Kormos & Trebits, 2012; Ong & Zhang, 2010; Révész et al., 2017; Tavakoli, 2014), with the prediction that a task with greater complexity at the level of idea conceptualization would allow writers to formulate more complex language (see Robinson, 2001, 2005).

Studies have employed picture-based writing tasks for conceptualization-level manipulations (e.g., Kormos, 2011; Kormos & Trebits, 2012; Tavakoli, 2014). Specifically, a picture narration task requires writers to develop their own story plot based on randomly ordered pictures, while a cartoon description task provides a clear storyline. Of the two tasks, the picture narration involving the writer's creation of a storyline is considered more cognitively complex. Examining these tasks, Kormos (2011) found that writers used more sophisticated lexical items for a more complex task but there was no difference between the two tasks in the linguistic complexity of their essays. Additionally, the degree of conceptual demands has been operationalized as the provision of supporting ideas for argumentation (e.g., Ong, 2014; Ong & Zhang, 2010; Révész et al., 2017). For example, Ong and Zhang (2010) explored the effect of idea support on language, and their findings showed significant effects of idea support on lexical diversity (more complex tasks eliciting increased lexical diversity) but little effect on fluency. Révész et al. (2017) also examined how idea support affects linguistic complexity (i.e., lexical sophistication, lexical diversity, and syntactic complexity) by analyzing argumentative essays composed by advanced L2 learners and found that the manipulation of idea support had a significant effect on lexical sophistication, but not on lexical diversity and syntactic complexity. These mixed findings of previous research might have arisen from some task and learner factor differences; Kormos (2011) and Révész et al. (2017) examined different genres, and although Ong and Zhang (2010) and Révész et al. commonly examined argumentative writing, they targeted L2 learners at different proficiency levels.

2.2. Genre as a task variable in L2 writing research

Writers are expected to communicate different functions in different genres. For example, narrative essays entail accounts of connected events or people's actions in a specific time frame, while non-narrative essays involve the argumentation or explanation of general ideas (Berman & Slobin, 1994). Genre studies have examined patterns of language variation across genres and explained writers' use of different linguistic features as an attempt to fulfill different rhetorical functions (e.g., Lu, 2011; Qin & Uccelli, 2016; Staples & Reppen, 2016; Yoon & Polio, 2017). Findings of genre studies have commonly indicated that non-narrative essays tend to feature syntactically more complex language than narrative essays. Lu (2011) examined the syntactic complexity of L2 narrative and

argumentative essays, and found that L2 learners used longer production units (e.g., mean sentence length and mean clause length) and more complex phrases (e.g., complex nominals and coordinate phrases) in argumentative than in narrative essays. This finding of different linguistic features across genres might indicate L2 writers' awareness of the need to express genre-specific functions using different linguistic resources.

Genre has also been operationalized as a task complexity variable (e.g., Alexopoulou et al., 2017; Ruiz-Funes, 2015; Yang, 2014). This line of research is based on the assumption that argumentative tasks that involve logical causal reasoning would be more cognitively demanding to L2 learners than narrative tasks. For example, Yang (2014) examined the four genres of narrative, expository, expo-argumentative and argumentative essays (in order of increasing cognitive complexity) composed by adult Chinese EFL learners. Her results showed the learners' use of simpler syntactic structures (unit-length and phrasal coordination measures) in the narrative essays than in the argumentative essays, but no significant effect of genre on fluency and accuracy. Yang interpreted her findings as evidence in partial support of Robinson's cognition hypothesis. Ruiz-Funes (2015) reported on two studies, each of which involved two writing genres. Her first study examined L2 learner performance on analytic and argumentative tasks, and the second study on personal narrative and expository tasks. In these studies, Ruiz-Funes operationalized the argumentative and expository tasks as more cognitively demanding than the analytic and narrative tasks. The results of these studies lacked statistical significance because of small sample sizes, but they indicated consistent patterns of increased syntactic complexity but decreased accuracy and fluency for the more complex tasks, arguably supporting Skehan's limited attentional capacity model.

Task-based studies have mainly focused on exploring whether their results correspond to the predictions of the task complexity frameworks. That is, a researcher manipulates a task variable based on the assumption of how it will affect the cognitive demands of the task imposed on learners and, as a result, their language use. Since learner perceptions of language tasks may not always match the intention of the researcher because of their different learning experiences, Révész (2014) called for the use of an independent measure for task perceptions. Returning to the discussion of genre as a task complexity variable, we can argue that L2 learners may not find it more cognitively demanding to compose argumentative than other writing tasks because of their extensive experience with argumentative writing as a typical genre for standardized testing and higher education (Qin & Karabacak, 2010; Wolfe, 2011). Some studies now employ a separate measure to confirm the actual effect of task manipulations on learner perceptions (e.g., Lee, 2020; Révész et al., 2017; Vasylets et al., 2017). Particularly, considering that linguistic features in writing reflect the functional and communicative needs of a particular task type (Frear & Bitchener, 2015; Staples & Reppen, 2016), we need to achieve a clearer understanding of what motivates written language differences.

Therefore, this study aims to explicate distinct roles of communicative functions and cognitive demands in shaping linguistic features. It examines argumentative and narrative writing (two functionally distinct genres widely employed in previous genre research), and targets the level of conceptual demands operationalized as the provision of supporting ideas, which is considered a valid writing task variable (Révész et al., 2017). A self-rating questionnaire is used to measure L2 learner perceptions of the writing tasks, and syntactic and lexical complexity measures at multiple levels are analyzed to explore the connection between learners' task perceptions and language use. This study addresses the following questions:

- 1 How do the manipulations of genre and idea support influence L2 learners' perceptions of writing tasks?
- 2 How do the manipulations of genre and idea support influence measures of linguistic complexity in L2 learners' writing?

3. Method

3.1. Participants

This study examined 304 essays composed by 76 ESL students at a high intermediate level. They received four to six hours of L2 writing instruction per week during a 15-week-long semester. The main objectives of the courses were to prepare students for university-level academic courses and to help them understand the importance of audience in academic writing. Their specific goals included developing students' summarizing, paragraphing, and revising skills, as well as guiding them in completing several multi-draft and timed writing tasks. The largest portion of the course grades involved multi-draft essay writing and revisions (50 %–70 % of the course grade). Much of the class time was used to develop various academic writing skills (e.g., building paragraphs with evidence), so students' participation in this research was their major practice for a timed writing exam at the end of the semester.

Forty-six male and 30 female undergraduate students participated, and their ages ranged from 18 to 27, with a mean of 19. They came from various countries. Fifty participants were Chinese native speakers, and seven were Arabic speakers. The remaining participants were native speakers of Korean ($n = 6$), Japanese ($n = 3$), Portuguese ($n = 3$), Malay ($n = 2$), Thai ($n = 2$), Turkish ($n = 2$), and French ($n = 1$). Their mean length of English study was approximately nine years, and their mean length of stay in the United States was 15 months, which indicates that the participants' English learning had been mostly in their native countries.

3.2. Instruments

Four writing prompts (two narrative and two argumentative prompts) were designed for this research. The narrative prompts involved narrating a personal story related to foreign language learning and use, while the argumentative prompts required participants to make logical arguments on very similar topics in an attempt to remove potential topic effects on language use. Task repetition effects were minimized by developing somewhat distinct prompts (e.g., narrative prompts eliciting positive or negative experiences, respectively). When developing the task prompts, I consulted three SLA experts and one test developer, and improved the quality of the

Table 1
Counterbalanced Data Collection Procedures.

Group	Session 1	Session 2	Session 3	Session 4
A	Nar/-Support	Arg/-Support	Nar/+Support	Arg/+Support
(n = 20)	Task questionnaire	Task questionnaire	Task questionnaire	Task questionnaire Background survey
B	Arg/+Support	Nar/-Support	Arg/-Support	Nar/+Support
(n = 19)	Task questionnaire	Task questionnaire	Task questionnaire	Task questionnaire Background survey
C	Nar/+Support	Arg/+Support	Nar/-Support	Arg/-Support
(n = 18)	Task questionnaire	Task questionnaire	Task questionnaire	Task questionnaire Background survey
D	Arg/-Support	Nar/+Support	Arg/+Support	Nar/-Support
(n = 19)	Task questionnaire	Task questionnaire	Task questionnaire	Task questionnaire Background survey

Note. Nar/-Support = narrative task with no idea support; Nar/+Support = narrative task with idea support; Arg/-Support = argumentative task with no idea support; Arg/+Support = argumentative task with idea support.

prompts based on their feedback (see Appendix A for the full prompts).

Within each genre, I manipulated the level of conceptual demands, operationalized as the provision of supporting ideas. The writing prompts with lower conceptual demands included some additional information (i.e., example storylines for narrative and supporting reasons for argumentative). I predicted that the ESL students would find the tasks with idea support cognitively less demanding and less difficult than the tasks with no such support. Throughout the manuscript, the narrative and argumentative prompts with idea support are labeled as *Nar/+Support* and *Arg/+Support* respectively (the prompts with no idea support as *Nar/-Support* and *Arg/-Support*).

Self-rating results using a task questionnaire have been shown to reflect the cognitive effects of task manipulations on learners validly (e.g., more complex tasks rated as cognitively more demanding). A self-rating questionnaire was given to the participants immediately after they completed each writing task. The questionnaire used for a larger project originally consisted of six 9-point Likert scale items (Révész et al., 2016), but the two items on task motivation and interest were excluded because they were less relevant to the aim of the current research. The remaining items assessed the participants' perceptions of each task in terms of (1) the mental effort induced by the task (i.e., task complexity), (2) task difficulty, (3) task anxiety, and (4) task confidence. Robinson (2001, 2007) assessed task perceptions with no inclusion of a separate item related to mental effort. However, Révész et al. (2016) found task-induced cognitive demands to be multidimensional. Following their suggestion, I differentiated between task complexity and task difficulty in the questionnaire so that their distinct constructs can be fully captured. The following are the task questionnaire items:

This task required no mental effort at all.	1 2 3 4 5 6 7 8 9	This task required extreme mental effort.
This task was not difficult at all.	1 2 3 4 5 6 7 8 9	This task was extremely difficult.
I felt really relaxed doing this task.	1 2 3 4 5 6 7 8 9	I felt frustrated doing this task.
I didn't do well on this task.	1 2 3 4 5 6 7 8 9	I did well on this task.

3.3. Procedure

The writing tasks were administered to students as part of the classroom curriculum. The students were informed that research participation was completely voluntary. To minimize potential testing effects from a repeated-measures design, data were collected at one-week intervals, with the order of the topics counterbalanced (see Table 1 for the summary of data collection procedures). I used only data from the participants who completed all four writing sessions.

For each of the writing sessions, participants composed timed essays (each under the time constraint of 30 min). They were not allowed to use dictionaries or other resources while writing. Immediately after writing, they were asked to complete a task perception questionnaire. In the last week of data collection, participants completed a survey on their background. Spelling errors were removed manually as text pre-processing for automated analyses. The average length of the essays was 301 words ($SD = 71$). The shortest essay included 152 words, while the longest comprised 489 words.

3.4. Linguistic complexity measures

I employed three natural language processing tools to take full account of the multidimensional nature of linguistic complexity (Lu, 2010; Norris & Ortega, 2009): L2 Syntactic Complexity Analyzer (SCA; Lu, 2010), Coh-Metrix (McNamara et al., 2014), and the Multidimensional Analysis Tagger (MAT; Nini, 2019). Given a very large number of linguistic features computed by these tools, I selected target measures based on the criteria of validity, construct distinctiveness, and redundancy. For example, the number of clauses per sentence (C/S) is a SCA measure of clausal embeddings that taps into both subordination and coordination, but these two constructs, given their reflection of different levels of language development (e.g., Bardovi-Harlig, 1992; Norris & Ortega, 2009), needed to be explored through distinct measures (e.g., clauses per T-unit (C/T) for subordination and T-units per sentence (T/S) for

Table 2
Targeted Clause Types and Examples.

Type	Example
Nominal clause	I discovered that each culture has its own communication method.
Adverbial clause	While I was talking, other people started to interrupt.
Adjectival clause	Individuals who can speak foreign language can spread their own culture to foreigners.

Table 3
Target Linguistic Complexity Measures.

Construct	Measure	Description	Tool
Length of unit	Mean length of sentence (MLS)	# of words / # of sentences	SCA
	Mean length of clause (MLC)	# of words / # of clauses	SCA
	Clauses per T-unit (C/T)	# of clauses / # of T-units	SCA
Subordination	Nominal clause density (NOMC)	# of nominal clauses * 1000 / # of words	MAT
	Adverbial clause density (ADVC)	# of adverbial clauses * 1000 / # of words	MAT
	Adjectival clause density (ADJC)	# of adjectival clauses * 1000 / # of words	MAT
	Coordinate phrases per clause (CP/C)	# of coordinate phrases / # of clause	SCA
	Complex nominals per clause (CN/C)	# of complex nominals / # of clauses	SCA
Phrase-level complexity	Left embeddedness (LEFT)	# of words before the main verb	Coh-Metrix
	Modifiers per noun phrase (Modifiers/NP)	# of modifiers / # of noun phrases	Coh-Metrix
Lexical features	vocd-D (D)	Based on vocd-D formula	Coh-Metrix
	Word frequency (WF)	Based on the CELEX corpus	Coh-Metrix

coordination). Some measures such as verb phrases per T-unit and complex T-units per T-unit were excluded because they had been found to be less valid as language development indicators (Lu, 2011). Of the three unit-length measures that SCA generates, I included the mean length of sentence (MLS) and the mean length of clause (MLC) because these two measures, unlike the mean length of T-unit, were shown to reflect two distinct constructs (clause-level and phrase-level complexity, respectively; Yoon, 2017). For other multiple measures tapping into a very similar construct (e.g., complex nominals per T-unit and complex nominals per clause), this study included only one measure that had the clause as its base unit.

Additionally, it includes three specific measures of subordination (nominal clauses, adverbial clauses, and adjectival clauses). A nominal clause serves as the object of a superordinate verb, with an optional complementizer. An adverbial clause modifies a superordinate verb, and its association with a main clause is through a subordinating conjunction. An adjectival clause, or a relative clause, modifies a noun to specify its meaning, with an optional relative pronoun (Collins & Hollo, 2010). Table 2 illustrates each of these clauses with examples from the dataset. These specific measures are included because research has shown that subordination as a unitary construct (e.g., overall subordination ratio) often fails to detect language development over a short period of time (e.g., Bulté & Housen, 2014; Mazgutova & Kormos, 2015) and is not sensitive enough to reflect genre variation (e.g., Lu, 2011; Yoon & Polio, 2017).

Normalized values of these subordination measures (occurrences per 1000 words) were calculated using MAT (Nini, 2019), which was built based on the Stanford Tagger (Toutanova et al., 2003). The nominal clause density involved frequencies of that verb complements, subordinator that deletion, and wh-clauses. The adverbial clause density included frequencies of past participial clauses, present participial clauses, causative adverbial subordinators, concessive adverbial subordinators, conditional adverbial subordinators, and other adverbial subordinators. The density of adjectival clauses was calculated using frequencies of “that” relative clauses (both subject and object positions), pied-piping relative clauses, wh-relative clauses (both subject and object positions), past participial relatives, and present participial relatives (see Nini, 2019).

This study also targets phrase-level complexity that has been found to predict language development and overall writing quality (e.g., Biber et al., 2011; Bulté & Housen, 2014; Crossley & McNamara, 2014; Lu, 2011). Specifically, it includes the number of complex nominals per clause, number of words before the main verb (degree of left embeddedness), modifiers per noun phrase, and coordinate phrases per clause, each of which involves a distinct aspect of noun phrase sophistication. It also includes two lexical measures (word frequency and vocd-D). Word frequency (WF) is a measure of lexical sophistication, which is a construct based on the assumption that more advanced writing would include a greater number of infrequent lexical items (Kyle & Crossley, 2014; Laufer & Nation, 1995). The Coh-Metrix WF measure is the average word frequency for all words extracted from the CELEX database, and lower WF values indicate greater lexical sophistication. Vocd-D (D) is a lexical diversity measure relatively insensitive to text length (McCarthy & Jarvis, 2010) that has been found to be a valid index of language development (Crossley et al., 2010; Lu, 2011). In terms of the validity of the automated processing tools, Polio and Yoon (2018) offered evidence that both SCA and Coh-Metrix are capable of capturing variation in syntactic complexity reliably (high correlations between hand coding and the automated tools on most measures) and validly (detection of significant genre differences). For the subordination measures from MAT, 20 % of the essay data ($n = 60$) were randomly selected and hand coded to check their reliability. The correlations between hand coding and the automated tool were found to be acceptable (nominal: $r = 76$; adverbial: $r = 67$; adjectival: $r = 88$). Table 3 presents a summary of the linguistic complexity measures included in this study.

Table 4
Multilevel Models of Task Perceptions, as Predicted by Genre and Idea Support.

Item			<i>b</i>	SE	<i>t</i>	<i>p</i>
Complexity	<i>Fixed effects</i>	(Intercept)	5.36	0.14	39.24	< .001***
		Genre	0.11	0.16	0.69	.490
		Idea Support	−0.41	0.16	−2.56	.011*
		Genre*Idea Support	0.14	0.32	0.45	.665
	<i>Random effects</i>	Subject	0.92			
		Residual	1.99			
Difficulty	<i>Fixed effects</i>	(Intercept)	4.91	0.15	33.81	< .001***
		Genre	0.17	0.17	1.00	.318
		Idea Support	−0.50	0.17	−2.92	.003**
		Genre*Idea Support	0.29	0.34	0.85	.398
	<i>Random effects</i>	Subject	1.05			
		Residual	2.22			
Anxiety	<i>Fixed effects</i>	(Intercept)	4.80	0.15	31.15	< .001***
		Genre	0.01	0.20	0.01	.999
		Idea Support	−0.33	0.20	−1.67	.097
		Genre*Idea Support	−0.05	0.39	−0.13	.894
	<i>Random effects</i>	Subject	1.06			
		Residual	2.96			
Confidence	<i>Fixed effects</i>	(Intercept)	5.11	0.13	39.93	< .001***
		Genre	0.19	0.19	1.01	.304
		Idea Support	−0.01	0.19	0.04	.972
		Genre*Idea Support	−0.72	0.37	−1.96	.052
	<i>Random effects</i>	Subject	0.60			
		Residual	2.60			

Note. Baselines for categorical variables: Genre = Argumentative; Idea Support = No Support.

*** $p < .001$.

** $p < .01$.

* $p < .05$.

4. Results

4.1. Task perceptions across task types

The descriptive results of the perception data are presented in Table B1 in Appendix B. To examine the effects of genre and idea support manipulations on perceptions statistically, multilevel analyses were computed using the lme4 package in R (Bates et al., 2015), with *genre* (argumentative and narrative) and *idea support* (no support and support) as fixed and *subject* as random effects. For multilevel modeling, deviation coding was used so that the intercept in each model can represent the grand mean rather than the baseline condition (i.e., coefficients interpreted directly as main effects). The result showed that the manipulation of idea support actually led to significant changes in the students' perceptions of task complexity and difficulty (see Table 4; complexity: estimate = −0.41, SE = 0.16, $t = -2.56$, $p = .011$; difficulty: estimate = −0.50, SE = 0.17, $t = -2.92$, $p = .003$). On the other hand, the students did not perceive different genres as imposing significantly different levels of complexity and difficulty (complexity: estimate = 0.11, SE = 0.16, $t = 0.69$, $p = .490$; difficulty: estimate = 0.17, SE = 0.17, $t = 1.00$, $p = .318$). These results might confirm the role of supporting ideas as a task variable affecting cognitive complexity in writing and refute the general assumption that argumentative tasks would be cognitively more demanding to L2 writers than narrative tasks. Additionally, task anxiety and task confidence did not show any significant interaction or main effects.

4.2. Linguistic complexity across task types

Tables 5 and 6 present the multilevel models that include linguistic complexity as the outcome variables (see Table B2 in Appendix B for descriptive results). The result showed a significant interaction between genre and idea support on MLS (estimate = −1.20, SE = 0.58, $t = -2.05$, $p < .041$). Stronger interaction effects were found on two phrase-level complexity measures (CN/C: estimate = −0.24, SE = 0.05, $t = -4.49$, $p < .001$; Modifiers/NP: estimate = −0.11, SE = 0.03, $t = -4.35$, $p < .001$). As visualized in Fig. 1, the result of post-hoc analyses (paired samples t -tests) indicated that the provision of idea support in narrative writing resulted in a significant decrease in noun phrase complexity (CN/C: $t(75) = 5.22$, $p < .001$, $d = 0.66$; Modifiers/NP: $t(75) = 2.95$, $p = .004$, $d = 0.33$), whereas a very similar manipulation in argumentative writing led to its significant increase (or an increasing pattern with no statistical significance) (CN/C: $t(75) = -1.46$, $p = .150$, $d = -0.14$; Modifiers/NP: $t(75) = -3.14$, $p = .002$, $d = -0.31$). The effect sizes for the post-hoc results were generally small (Plonsky & Oswald, 2014).

In contrast to the limited effects of manipulating idea support on language use, main effects of genre were prevalent for many linguistic measures. Specifically, the argumentative essays elicited significantly higher values of unit length (MLS: estimate = −1.56, SE = 0.29, $t = -5.37$, $p < .001$; MLC: estimate = −1.28, SE = 0.14, $t = -9.28$, $p < .001$) and phrase-level complexity (CP/C: estimate = −0.03, SE = 0.02, $t = -2.88$, $p = .004$; CN/C: estimate = −0.45, SE = 0.03, $t = -16.39$, $p < .001$; Left embedded: estimate = −0.92, SE

Table 5

Multilevel Models of Length of Unit and Subordination Measures, as Predicted by Genre and Idea Support.

Measure			<i>b</i>	SE	<i>t</i>	<i>p</i>
MLS	<i>Fixed effects</i>	(Intercept)	16.88	0.42	39.78	< .001***
		Genre	−1.56	0.29	−5.37	< .001***
		Idea Support	0.40	0.29	1.37	.173
		Genre* Idea Support	−1.20	0.58	−2.05	.041*
	<i>Random effects</i>	Subject	12.08			
		Residual	6.44			
MLC	<i>Fixed effects</i>	(Intercept)	9.12	0.13	74.48	< .001***
		Genre	−1.28	0.14	−9.28	< .001***
		Idea Support	−0.30	0.14	−2.17	.030*
		Genre* Idea Support	−0.51	0.28	−1.86	.065
	<i>Random effects</i>	Subject	0.78			
		Residual	1.46			
C/T	<i>Fixed effects</i>	(Intercept)	1.58	0.04	43.62	< .001***
		Genre	0.01	0.03	0.25	.803
		Idea Support	0.09	0.03	3.13	.002**
		Genre* Idea Support	−0.01	0.06	−0.09	.930
	<i>Random effects</i>	Subject	0.04			
		Residual	0.06			
NOMC	<i>Fixed effects</i>	(Intercept)	8.87	0.41	21.45	< .001***
		Genre	2.13	0.65	3.25	.001**
		Idea Support	−1.08	0.65	−1.65	.100
		Genre* Idea Support	0.81	1.31	0.62	.535
	<i>Random effects</i>	Subject	4.85			
		Residual	32.59			
ADVC	<i>Fixed effects</i>	(Intercept)	14.87	0.44	33.71	< .001***
		Genre	−3.30	0.85	−3.88	< .001***
		Idea Support	−0.73	0.85	−0.86	.393
		Genre* Idea Support	−0.83	1.70	−0.49	.624
	<i>Random effects</i>	Subject	10.81			
		Residual	31.03			
ADJC	<i>Fixed effects</i>	(Intercept)	9.82	0.49	19.87	< .001***
		Genre	−2.69	0.64	−4.21	< .001***
		Idea Support	1.02	0.64	1.59	.113
		Genre* Idea Support	−1.37	1.28	−1.07	.285
	<i>Random effects</i>	Subject	10.81			
		Residual	31.03			

Note. Baselines for categorical variables: Genre = Argumentative; Idea Support = No Support; MLS = mean length of sentence; MLC = mean length of clause; C/T = clauses per T-unit; NOMC = nominal clause density; ADVC = adverbial clause density; ADJC = adjectival clause density.

*** $p < .001$.

** $p < .01$.

* $p < .05$.

= 0.15, $t = -6.20$, $p < .001$; Modifiers/NP: estimate = −0.18, SE = 0.01, $t = -14.59$, $p < .001$) than the narrative essays. The clauses per T-unit (C/T) measure, which had been extensively adopted as a typical measure of clausal subordination, was not shown to change across the two genres, and this result is in line with the findings of previous research (e.g., Lu, 2011; Yoon & Polio, 2017). However, using more fine-grained measures of clausal subordination (i.e., nominal, adverbial, and adjectival clause density), I found that narrative writing is characterized by increased nominal clause density (estimate = 2.13, SE = 0.65, $t = 3.25$, $p = .001$), while argumentative writing by increased density of adverbial clauses (estimate = −3.30, SE = 0.85, $t = -3.88$, $p < .001$) and adjectival clauses (estimate = −2.69, SE = 0.64, $t = -4.21$, $p < .001$). This result is notable in that, unlike previous studies attending to phrasal measures and, accordingly, rejecting clausal subordination in relation to genre variation (except for Frear & Bitchener, 2015; Staples & Reppen, 2016), the result clearly indicates that the use of more specific measures allows us to detect how different genres elicit different characteristics of clausal subordination in L2 writing (see Fig. 2), which has been gone unnoticed in previous research due to its reliance on a general subordination measure.

Table 7 presents a summary of the statistical analyses for task perceptions and linguistic features, showing that textual changes across task types had little to do with the learners' perceptions of the tasks. Previous studies have tested the validity of task complexity (i.e., task manipulation effects on cognitive task complexity) with regard to significant changes in linguistic features, mostly tapping into the construct of linguistic complexity. However, this study found that the provision of idea support actually resulted in a significant decrease in L2 learners' task complexity and difficulty, but this valid manipulation of task complexity did not push the learners to use different linguistic resources for task completion, challenging a widely accepted hypothesis in task-based research.

Conversely, genre variation, which was shown to have little influence on L2 learners' perceptions of task complexity and difficulty, led them to use different language for writing. This finding may suggest the need to disentangle the effects of task manipulation on learner perceptions from those on language production because the different cognitive demands of writing tasks do not necessarily result in different linguistic structure use, potentially due to the characteristics of writing that involve planning and revision processes

Table 6
Multilevel Models of Phrasal Complexity and Lexical Complexity Measures, as Predicted by Genre and Idea Support.

Measure			<i>b</i>	SE	<i>t</i>	<i>p</i>
CP/C	<i>Fixed effects</i>	(Intercept)	0.18	0.01	20.70	< .001***
		Genre	−0.03	0.02	−2.88	.004**
		Idea Support	−0.01	0.02	−1.04	.301
		Genre* <i>Idea Support</i>	−0.02	0.02	−0.65	.518
	<i>Random effects</i>	Subject	0.01			
		Residual	0.01			
CN/C	<i>Fixed effects</i>	(Intercept)	1.08	0.03	42.77	< .001***
		Genre	−0.45	0.03	−16.39	< .001***
		Idea Support	−0.06	0.03	−2.38	.018*
		Genre* <i>Idea Support</i>	−0.24	0.05	−4.49	< .001***
	<i>Random effects</i>	Subject	0.03			
		Residual	0.06			
LEFT	<i>Fixed effects</i>	(Intercept)	4.53	0.12	38.25	< .001***
		Genre	−0.92	0.15	−6.20	< .001***
		Idea Support	−0.14	0.15	−0.91	.364
		Genre* <i>Idea Support</i>	−0.05	0.30	−0.18	.855
	<i>Random effects</i>	Subject	0.64			
		Residual	1.69			
Modifiers/NP	<i>Fixed effects</i>	(Intercept)	0.70	0.01	67.45	< .001***
		Genre	−0.18	0.01	−14.59	< .001***
		Idea Support	0.01	0.01	0.99	0.321
		Genre* <i>Idea Support</i>	−0.11	0.03	−4.35	< .001***
	<i>Random effects</i>	Subject	0.01			
		Residual	0.01			
D	<i>Fixed effects</i>	(Intercept)	78.16	1.28	61.10	< .001***
		Genre	4.39	1.45	3.02	.003**
		Idea Support	2.82	1.45	1.94	.054
		Genre* <i>Idea Support</i>	3.23	2.91	1.11	.268
	<i>Random effects</i>	Subject	84.25			
		Residual	160.47			
WF	<i>Fixed effects</i>	(Intercept)	3.08	0.01	441.81	< .001***
		Genre	0.04	0.01	6.37	< .001***
		Idea Support	0.02	0.01	3.62	< .001***
		Genre* <i>Idea Support</i>	−0.02	0.02	−1.74	.083
	<i>Random effects</i>	Subject	0.01			
		Residual	0.01			

Note. Baselines for categorical variables: Genre = Argumentative; Idea Support = No Support; CP/C = coordinate phrases per clause; CN/C = complex nominals per clause; LEFT = left embeddedness; Modifiers/NP = modifiers per noun phrase; D = vocd- D; WF = word frequency.

*** $p < .001$.

** $p < .01$.

* $p < .05$.

(Hayes, 1996, 2012). That is, the cognitive burden of a writing task may not be a major factor influencing the formulation of linguistic structures, as writers can spend much time in generating ideas and selecting syntactic and lexical units through planning. They are also allowed to read and revise their text after execution. The association between task complexity and linguistic complexity was further explored by conducting Pearson correlations between perceived task complexity and linguistic complexity features for each of the task types. As Table 8 presents, there were indeed limited relationships between task complexity, approximated by L2 learners' perceptions of the tasks, and their use of language.

5. Discussion

5.1. Perceptions of writing tasks

It has been widely accepted that argumentative tasks would be perceived more cognitively demanding than narrative tasks because argumentation necessitates learners to use higher-order reasoning and interpretations that go beyond experience and knowledge telling for narration (Bereiter & Scardamalia, 1987). However, the results of this study indicated that L2 learners view both genres as imposing similar cognitive demands, and this can possibly be explained by the prediction of cognitive models of writing (Hayes, 1996, 2012) that cognitive pressures are influenced by genre schemas, task schemas, and other long-term memory factors. That is, many adult learners who have experienced argumentative tasks for educational and testing purposes possess well-established genre schemas for argumentation, leading to a reduced processing burden during argumentative writing despite the inherent, higher-level cognitive loads of this genre.

Although there was little effect of genre on task complexity and difficulty, this study found that the manipulation of idea support significantly influenced the learners' perceptions of task complexity and difficulty. This finding confirms the results of Révész et al.

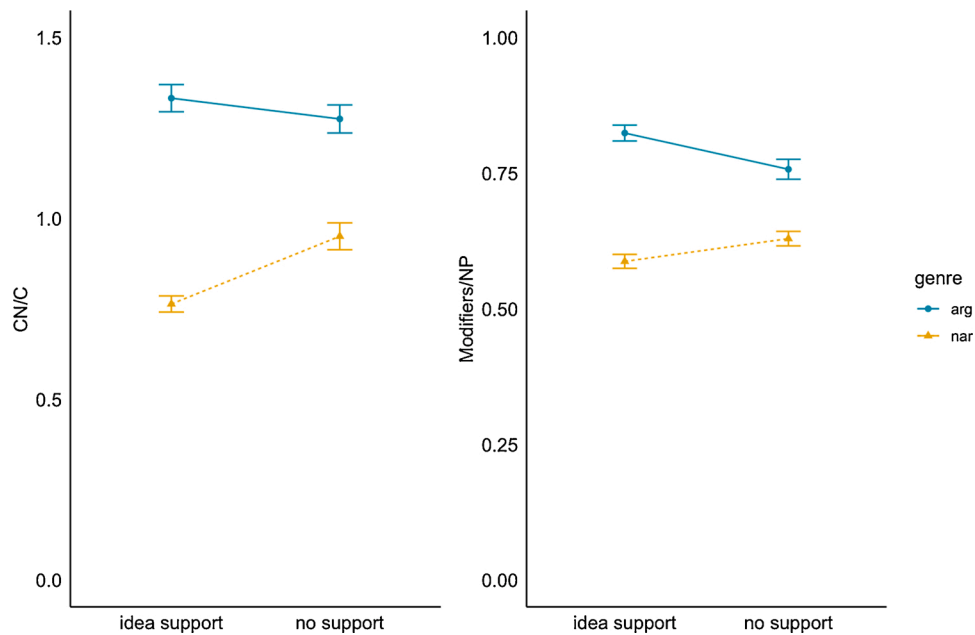


Fig. 1. Interaction plots for complex nominals per clause and modifiers per noun phrase showing an interaction between genre and idea support conditions.

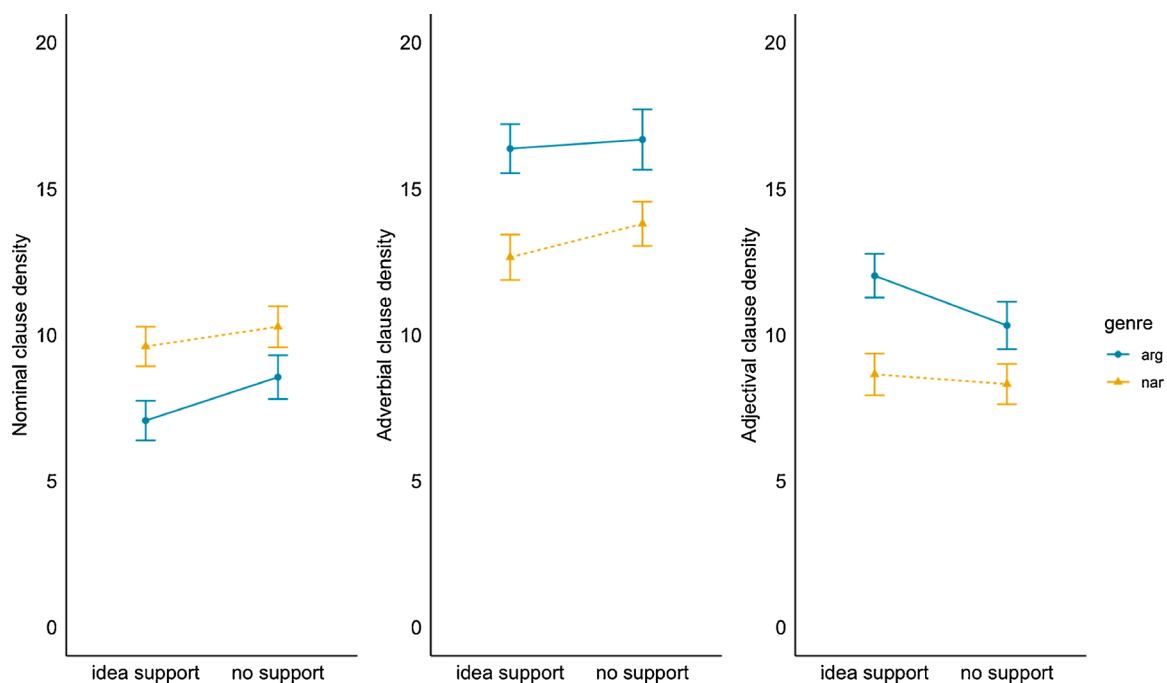


Fig. 2. Nominal clause density, adverbial clause density, and adjectival clause density across genre and idea support conditions.

(2017) from argumentative tasks and further indicates that idea support may be a valid task variable not only in argumentative but also in other genres. Taking these findings as a starting point, further studies can be conducted to explore how to maximize the intended effect of idea support manipulations in various writing tasks and test the applicability of other task elements to written discourse (e.g., elements in the triadic componential framework; Robinson, 2007).

Table 7

Summary of Task Manipulation Conditions with Significant Results.

Measures	Genre	Idea Support	Genre*Idea Support
Task perceptions			
Task complexity	–	-Support > +Support	–
Task difficulty	–	-Support > +Support	–
Task anxiety	–	–	–
Task confidence	–	–	–
Linguistic complexity			
Length of production	Argument > Narrative	–	–
Clauses per T-unit	–	+Support > -Support	–
Nominal clause density	Narrative > Argument	–	–
Adverbial clause density	Argument > Narrative	–	–
Adjectival clause density	Argument > Narrative	–	–
Noun phrase complexity	Argument > Narrative	–	+Support: Higher in Arg and Lower in Nar
Lexical diversity	Narrative > Argument	–	–
Lexical sophistication	Argument > Narrative	-Support > +Support	–

Note. The greater than symbol (>) indicates that the first condition is significantly greater than the second. For example, “Argument > Narrative” for noun phrase complexity means more complex noun phrases in argumentative than in narrative essays.

Table 8

Correlations of Perceived Task Complexity with Linguistic Complexity Features.

Linguistic features	Task complexity			
	Arg/-Support	Arg/+Support	Nar/-Support	Nar2/+Support
MLS	.217	.090	-.205	-.103
MLC	.146	.177	-.221	-.046
C/T	.072	.009	-.099	-.062
NOMC	-.081	-.051	-.096	.113
ADVC	.120	-.064	-.174	.057
ADJC	.081	-.126	-.098	-.079
CP/C	.064	.030	-.162	-.077
CN/C	.208	.107	-.189	.013
LEFT	.123	.093	-.269*	.217
Modifiers/NP	.193	.149	-.167	.016
D	-.212	-.018	-.168	.077
WF	.063	-.176	.106	-.171

Note. *correlations are significant at the alpha level of .05.

5.2. Effects of task manipulations on linguistic complexity features

This study found that learners' use of syntactic structures differed widely across the two genres, whereas their language differed to a limited extent across the two idea support conditions. Of the targeted linguistic measures, nominal complexity (complex nominals per clause and modifiers per noun phrase) was found to differ to the largest extent across the two genres. Below are excerpts from the essays written by one student writer. In the excerpts, the instances of complex nominals are underlined.

You will be provided greater job opportunities related to international business. This opportunity is valuable since there are huge markets in other countries. Those who can speak many languages have earned a lot of money from international business. Moreover, by having a good command of a foreign language, you gain more fun from various activities such as traveling or watch foreign TV programs. You can enjoy different kind of view and broaden your horizons. This is a very cool experience that definitely worth a try. (Arg/+Support, Student 45)

Everything went well at first, and I was quite satisfied with my new circumstance here. The air was clean and fresh, and the sky was pure blue. I can seldom enjoy this kind of environment in my hometown. I was in good mood, and well prepared to start my study life here, until that day I went to my first Mathematics class. I found my classroom easily and took a seat there. I was nervous since I was unfamiliar with the American teaching style. (Nar/+Support, Student 45)

As shown in the excerpts, argumentative texts tend to include more complex nominals than narrative texts, indicating that the use of complex noun phrases concerns an issue beyond language development but rather relates to the selection of appropriate linguistic resources in different rhetorical situations.

Below are example excerpts intended to show how various types of dependent clauses appear differently in argumentative and narrative essays (nominal clauses in bold, adverbial clauses double underlined, and adjectival clauses underlined). In interpreting these excerpts, I focus only on nominal and adjectival clauses since they have fairly contrastive functions.

... increasingly amount of countries are seeking the opportunities of cooperating with China, so the people who have the ability to speak other languages have more chances to participate in international events. In the meantime, the rise of international companies gives

people more job opportunities, and most of the jobs they provide a relatively high income... On the other hand, you travel experience can be fantastic if you can understand the language that the country use. (Arg/+Support, Student 47)

*I remember **that I tried to ask somebody for the right path by using English**, because my friend said it's okay to say English to them, they'll understand. But soon I found out **that my biggest issue is not speaking correct English to them**, but I can't understand **what they reply in English**. Then I had to read their gesture, and a nice lady even used electronic dictionary in her phone to translate her word into English. Fortunately, most of them can understand **what you said in English**.* (Nar/+Support, Student 47)

An examination of the occurrences of nominal clauses revealed that narrative writing tends to include many *stative mental verb + nominal clause* constructions (verbs including *find out*, *remember*, and *understand*), whereas the argumentative excerpt does not contain any nominal clause (only one case from the entire essay). Given the major function of mental verbs for describing states and actions experienced by humans, this finding of increased nominal clauses in narrative writing may be evidence of L2 learners' attempt to describe their experience in an accurate way.

Supporting the findings of previous research (e.g., Lu, 2011; Qin & Uccelli, 2016), the results may indicate that genre indeed functions as a task variable eliciting different language use from L2 learners; argumentative tasks encourage learners to use more syntactically complex language than narrative tasks. Additionally, the result of the three fine-grained measures of subordination showed that the argumentative essays tended to contain more adverbial and adjectival clauses than the narrative essays did, but nominal clauses were more frequent in the narrative than the argumentative essays. This finding points to the importance of employing more specific measures when exploring genre or task type effects on language (Frear & Bitchener, 2015), as it contrasts with the findings of previous research (with only general subordination measures such as clauses per T-unit) showing little effect of genre on subordination (e.g., Lu, 2011; Yoon & Polio, 2017).

Compared to genre effects, idea support exerted much less influence on learners' language use. This finding may offer evidence of a limited relationship between language and cognitive complexity in writing, but previous studies and the current finding have commonly shown that lexical complexity is an area of language that changes with the level of cognitive complexity (Kormos, 2011; Ong & Zhang, 2010; Révész et al., 2017). It would be meaningful to further clarify the association between participants' perceptions of cognitive complexity and lexical features in their writing. In terms of syntactic complexity, the findings of this study indicated that intermediate-level L2 learners had sufficient awareness of the need to use different syntactic structures for different purposes, achieving genre-specific rhetorical functions. The significant interaction result on noun phrase complexity (CN/C and Modifiers/NP) suggests greater nominal complexity in the cognitively less demanding argumentative task but lower nominal complexity in the cognitively less demanding narrative task. It may also be evidence of the writer's endeavor to fulfill different communicative requirements using increased cognitive resources available for writing. Specifically, they use complex noun phrases to a greater extent to build more convincing arguments (e.g., relative clauses, attributive adjectives) in argumentative writing, while using fewer complex noun phrases in narrative writing because this genre requires more use of personal pronouns and place and time adverbials. Berman and Slobin (1994) noted that "the development of grammar cannot be profitably considered without attention to the psycholinguistic and communicative demands of the production of connective discourse" (p. 2). These results suggest that it would be necessary to further identify language structures that fulfill genre-specific communicative functions and also to use tasks that facilitate such structures effectively.

6. Conclusion

It has been argued that the perceived cognitive demands imposed by tasks influence learners' language use, and similarly previous findings of cross-genre variation in language have been explained as the outcome of different cognitive demands imposed by different genres. However, as evidenced by the result of this study, the cognitive complexity of the writing tasks does not relate strongly to the language elicited by the tasks. A possible explanation is that syntactic structures in writing are a manifestation of a set of communicative functions demanded by a particular genre. Thus, task-based writing research needs to be conducted with the awareness that there may be a stronger connection between genre and language than between cognitive complexity and language.

Nevertheless, the findings of this study need to be accepted with caution because it has some limitations. First, the data were collected from high-intermediate L2 learners, and their language use and genre awareness may not reflect how L2 learners would generally perform the same tasks. Also, this study employed timed writing tasks that elicited short essays. While the use of timed tasks was a proper methodological decision in this cognitively oriented task study, we note that, in instructional settings, timed writing needs to be supplemented with other writing tasks that allow students to build more extended and contextualized discourse. Regarding the writing prompts, this study employed one prompt for each of the task conditions, and the use of similar topics that aimed to remove topic effects might have potentially led to task repetition effects. The idea support prompts included some additional input with complex phrases, which might have affected the language use of the learners.

Pedagogical implications of this study involve how teachers need to understand and implement different genres for L2 writing instruction. Teachers are advised to have a clear awareness of their students' difficulties, needs, or preferences for different genres and what linguistic expressions are needed to accomplish the expected functions of targeted genres. Based on this understanding, writing tasks appropriate for the students can be designed, selected, and sequenced. Moreover, L2 writing instruction tends to focus on learners' improvement in argumentative writing skills and pay relatively less attention to other genres, due to the belief that argumentative writing is most appropriate for testing purposes. Teachers may also assume that their students do not need any more instruction for narrative writing when they show sufficient skills for argumentative writing because of their view of narrative writing as a simpler task than argumentative writing. However, the cultural and affective benefits of narrative tasks point to the need for greater attention to narrative tasks in L2 writing classes (e.g., Berman & Slobin, 1994; Zhang, 2013). Building a writing curriculum informed

by task and genre research would enable L2 learners to develop linguistic competence and genre awareness more effectively.

Taking an exploratory approach to investigating genre and task effects, this study showed the importance of interpreting written syntactic features as the linguistic realization of genre-specific communicative functions. It also offered evidence for the role of supporting ideas in reducing cognitive pressure. However, there is still an unresolved issue of how specific such supporting ideas should be in a prompt. As a next step, future research could examine the influence of different levels of specificity and amount for idea support on learners' task perceptions and language use. Replication research with learners at a lower proficiency level would help achieve a fuller picture of writing genre and task manipulation effects. A study with an increased number of writing prompts (e.g., multiple prompts for one manipulated task condition) and task conditions can be conducted to explore the validity of the current findings. Lastly, given a close link between rhetorical functions and syntactic features in writing, L2 writing research may need to focus more on how learners attempt to communicate their meaning through various syntactic structures than on how the cognitive demands of tasks push learners to use more complex syntax in written discourse.

Acknowledgements

I would like to acknowledge that the funding for this research has been provided by the International Research Foundation (TIRF Doctoral Dissertation Grant, 2016) and the National Federation of Modern Language Teachers Association (NFMLTA/MLJ Dissertation Support Grant, 2016)

Appendix A

Writing Prompts

Argumentative 1 (Arg/-Support)

Situation:

You attended a seminar and the main theme was that using a foreign language fluently has become necessary in this globalized era.

Writing task:

Write an essay about whether you agree or disagree with the statement about the necessity of foreign language abilities. Support your position with reasons. Be sure to fully develop your essay by including clear explanations and logical supporting ideas.

Argumentative 2 (Arg/+Support)

Situation:

You attended a seminar and the main theme was that the ability to speak a foreign language raises the possibility of having a successful life.

Writing task:

Write an essay about whether you agree or disagree with the statement about the relationship between foreign language abilities and success. Support your position using the reasons provided below. Be sure to fully develop your essay by including clear explanations and logical supporting ideas.

Agree/Support to argue for the position

- Better understanding of cultural differences and other ethnic groups
- Greater job opportunities related to international business
- Possibilities for fun activities such as traveling or watching foreign TV programs

Disagree/Support to argue against the position

- Other qualities (such as self-confidence) more important than foreign language skills
- Foreign language skills not necessary for many great jobs
- A huge investment of time and effort for language learning that could be used for other skill development

Narrative1 (Nar/-Support)

Situation:

Your friend has plans to learn a foreign language but is afraid it might be useless to spend the time learning a language. You have successfully learned a foreign language and use it often. You want to show your friend that language learning and use can be interesting by telling him/her about your positive experience.

Writing task:

Tell a story about ONE of your positive experiences related to foreign language use. Be sure to fully develop your story by including specific details.

Narrative 2 (Nar/+Support)

Situation:

Your friend is planning a trip to a foreign country. While excited about this trip, your friend is worried about how to communicate with people using a foreign language. You have greater foreign language experience, so your friend wants to know some of the possible difficulties she may have while interacting with foreigners.

Writing task:

Tell a story about ONE of your difficult experiences related to interactions using a foreign language. When developing your ideas, you can refer to the storylines below and use any of them to facilitate your writing. Be sure to fully develop your story by including specific details.

Example storylines

- You visited a public place in a foreign country. When you were talking to a foreigner, he/she corrected your language constantly, making you feel offended. Then...
- You were talking to a foreigner. While interacting with him/her, you experienced some cultural differences that made you feel uncomfortable. Then...
- You had to fix a problem or sign a contract using a foreign language. For such purposes, you expressed your ideas to a native speaker of the language, but it caused a misunderstanding, leading to a serious accident. Then...

Appendix B

Table B1

Descriptive Statistics for ESL Students' Perceptions of Writing Tasks.

Item	Arg/-Support		Arg/+Support		Nar/-Support		Nar/+Support	
	M (SD)	95 % CI	M (SD)	95 % CI	M (SD)	95 % CI	M (SD)	95 % CI
Complexity	5.55 (1.77)	[5.15, 5.96]	5.07 (1.73)	[4.67, 5.46]	5.60 (1.68)	[5.21, 5.97]	5.25 (1.65)	[4.87, 5.63]
Difficulty	5.14 (1.76)	[4.74, 5.55]	4.50 (1.66)	[4.12, 4.88]	5.17 (1.84)	[4.75, 5.59]	4.82 (1.96)	[4.24, 5.56]
Anxiety	4.95 (2.00)	[4.49, 5.40]	4.64 (2.04)	[4.18, 5.11]	4.97 (2.14)	[4.49, 5.46]	4.62 (1.83)	[4.20, 5.04]
Confidence	4.84 (1.86)	[4.42, 5.27]	5.20 (1.74)	[4.80, 5.59]	5.39 (1.76)	[4.99, 5.80]	5.03 (1.80)	[4.61, 5.44]

Table B2

Descriptive Statistics for Linguistic Complexity Measures by Task Type.

Measure	Arg/-Support		Arg/+Support		Nar/-Support		Nar/+Support	
	M (SD)	95 % CI	M (SD)	95 % CI	M (SD)	95 % CI	M (SD)	95 % CI
Length of unit								
MLS	17.17 (4.03)	[16.24, 18.09]	18.16 (4.52)	[17.13, 19.19]	16.20 (4.09)	[15.27, 17.14]	16.00 (4.54)	[14.96, 17.04]
MLC	9.79 (1.67)	[9.41, 10.17]	9.74 (1.38)	[9.42, 10.06]	8.76 (1.57)	[8.40, 9.12]	8.20 (1.34)	[7.90, 8.51]
Subordination								
C/T	1.58 (0.28)	[1.51, 1.64]	1.67 (0.34)	[1.59, 1.74]	1.59 (0.30)	[1.52, 1.65]	1.67 (0.34)	[1.59, 1.75]
NOMC	8.55 (6.52)	[7.06, 10.04]	7.06 (5.92)	[5.71, 8.42]	10.27 (6.11)	[8.88, 11.67]	9.60 (5.90)	[8.25, 10.95]
ADVC	16.68 (8.99)	[14.62, 18.73]	16.36 (7.31)	[14.69, 18.04]	13.79 (6.60)	[12.28, 15.30]	12.65 (6.77)	[11.10, 14.20]
ADJC	10.32 (7.08)	[8.70, 11.93]	12.02 (6.52)	[10.53, 13.51]	8.31 (6.02)	[6.94, 9.69]	8.64 (6.21)	[7.23, 10.06]
Phrasal complexity								
CP/C	0.20 (0.14)	[0.17, 0.23]	0.19 (0.11)	[0.17, 0.22]	0.17 (0.12)	[0.14, 0.20]	0.15 (0.09)	[0.13, 0.17]
CN/C	1.28 (0.34)	[1.20, 1.35]	1.33 (0.33)	[1.26, 1.41]	0.95 (0.32)	[0.88, 1.02]	0.76 (0.20)	[0.72, 0.81]
Left embedded	5.05 (1.82)	[4.63, 5.47]	4.94 (1.77)	[4.53, 5.35]	4.15 (1.12)	[3.90, 4.41]	3.99 (1.27)	[3.70, 4.28]
Modifiers/NP	0.76 (0.16)	[0.72, 0.79]	0.82 (0.13)	[0.79, 0.85]	0.63 (0.12)	[0.60, 0.66]	0.59 (0.11)	[0.56, 0.61]
Lexical complexity								
D	75.36 (13.84)	[72.20, 78.52]	76.57 (16.35)	[72.83, 80.30]	78.14 (17.07)	[74.24, 82.04]	82.57 (15.11)	[79.12, 86.02]
WF	3.04 (0.09)	[3.02, 3.06]	3.08 (0.08)	[3.06, 3.09]	3.09 (0.07)	[3.08, 3.11]	3.11 (0.08)	[3.09, 3.12]

References

- Alexopoulou, T., Michel, M., Murakami, A., & Meurers, D. (2017). Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques. *Language Learning*, 67, 180–208.
- Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26, 390–395.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *Package 'lme4'*. Retrieved from <https://cran.r-project.org/web/packages/lme4/lme4.pdf>.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Lawrence Erlbaum.
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Erlbaum.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45, 5–35.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65.
- Byrnes, H., & Manchón, R. M. (Eds.). (2014). *Task-based language learning—Insights from and for L2 writing* (Vol. 7). John Benjamins.

- Collins, P., & Hollo, C. (2010). *English grammar: An introduction*. Palgrave Macmillan.
- Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, 26, 66–79.
- Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2010). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing*, 28, 561–580.
- Frear, M. W., & Bitchener, J. (2015). The effects of cognitive task complexity on writing complexity. *Journal of Second Language Writing*, 30, 45–57.
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy, & S. Randall (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Erlbaum.
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29, 369–388.
- Housen, A., De Clercq, B., Kuiken, F., & Vedder, I. (2019). Multiple approaches to complexity in second language research. *Second Language Research*, 35, 3–21.
- Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing*, 37, 13–38.
- Johnson, M. D., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing*, 21, 264–282.
- Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing*, 20, 148–161.
- Kormos, J., & Trebits, A. (2012). The role of task complexity, modality, and aptitude in narrative task performance. *Language Learning*, 62, 439–472.
- Kyle, K., & Crossley, S. A. (2014). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly*, 49, 757–786.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307–332.
- Lee, J. (2020). Task closure and task complexity effects on L2 writing performance. *Journal of Second Language Writing*, 50, Article 100777.
- Lu, X. (2010). Automatic measurement of syntactic complexity in child language acquisition. *International Journal of Corpus Linguistics*, 14, 3–28.
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45, 36–62.
- Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3–15.
- McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42, 381–392.
- McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.
- Nini, A. (2019). The multi-dimensional analysis tagger. In T. B. Sardinha, & M. V. Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 67–94). Bloomsbury: Academic.
- Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, 555–578.
- Ong, J. (2014). How do planning time and task conditions affect metacognitive processes of L2 writers? *Journal of Second Language Writing*, 23, 17–30.
- Ong, J., & Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, 19, 218–233.
- Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912.
- Polio, C., & Yoon, H. (2018). The reliability and validity of automated tools for examining variation in syntactic complexity across genres. *International Journal of Applied Linguistics*, 28, 165–188.
- Qin, J., & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System*, 38, 444–456.
- Qin, W., & Uccelli, P. (2016). Same language, different functions: A cross-genre analysis of Chinese EFL learners' writing performance. *Journal of Second Language Writing*, 33, 3–17.
- Révész, A. (2014). Towards a fuller assessment of cognitive models of task-based learning: Investigating task-generated cognitive demands and processes. *Applied Linguistics*, 35, 87–92.
- Révész, A., Kourtali, N., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, 67, 208–241.
- Révész, A., Michel, M., & Gilabert, R. (2016). Measuring cognitive task demands using dual-task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies in Second Language Acquisition*, 38, 703–737.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27–57.
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics*, 43, 1–32.
- Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics*, 45, 193–213.
- Ruiz-Funes, M. (2015). Exploring the potential of second/foreign language writing for language learning: The effects of task factors and learner variables. *Journal of Second Language Writing*, 28, 1–19.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, 32, 17–35.
- Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes, & R. M. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 217–236). John Benjamins.
- Toutanova, K., Klein, D., Manning, C., & Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. *Proceedings of HLT-NAACL 2003*, 252–259.
- Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67, 394–430.
- Wolfe, C. R. (2011). Argumentation across the curriculum. *Written Communication*, 28, 193–219.
- Yang, W. (2014). *Mapping the relationships among the cognitive complexity of independent writing tasks, L2 writing quality, and complexity, accuracy and fluency of L2 writing*. Doctoral dissertation.
- Yoon, H. (2017). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System*, 66, 130–141.
- Yoon, H., & Polio, C. (2017). ESL students' linguistic development in two written genres. *TESOL Quarterly*, 51, 275–301.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24, 1–27.
- Zhang, L. J. (2013). Second language writing as and for second language learning. *Journal of Second Language Writing*, 22, 446–447.

Hyung-Jo Yoon is an Assistant Professor in the Department of Linguistics/TESL at California State University, Northridge. His research interests include second language writing, task-based language teaching, and language assessment, with a particular focus on building effective pedagogy for second language development. He also conducts research related to corpus linguistics and computational text analysis.