Structural Equation Modelling of EFL Learners' Perceived Preferences for Data-driven Learning and Learners' Agency

Jinfang Liang[1], & Ying Tan[1]

[1] International College, Krirk University, Bangkok 10220, Thailand

Correspondence: Jinfang Liang, International College, Krirk University, Bangkok 10220, Thailand

Received: January 5, 2023 Accepted: February 10, 2023 Online Published: February 14, 2023

doi:10.5430/wjel.v13n3p90 URL: https://doi.org/10.5430/wjel.v13n3p90

Abstract

Data-driven learning (DDL) has drawn researchers' eyes on corpus linguistics and language learning successfully, particularly on English writing. However, the structural relation between the students' preferences for data-driven learning and the EFL students' learning agency has not been well examined yet. This study examined the hypothetical model of measurement for EFL learners' perceived preferences for DDL and their learning agency. Two questionnaires were used for collecting the data. Structural equation modeling (SEM) was assessed using AMOS. The results revealed that the developed model enjoyed an acceptable level of goodness of fit. The results also showed that the students' perceived preferences for DDL strongly affect their learning agency. Therefore, it could be concluded that exposure to DDL fosters language learners' self-efficacy and the ability to self-regulate their learning activities. All in all, the results have implications (theoretical and practical) for language teachers, learners and those interested in corpus linguistics.

Keywords: data-driven learning, learners' agency, EFL learners, corpus linguistics

1. Introduction

The term DDL, initially introduced by Johns (1991), explains how language learners can independently examine the language and detect the underlying regularities and rules inductively for themselves. The learner is regarded as a detective whose task is to explore the linguistic rules compiled from the corpus, which are conceptualized as "an electronically stored, searchable collection of texts" (Jones & Waller, 2015, p. 5). The key mechanism for this approach is that learners are assigned an active role in discovering the structural patterns for themselves following being exposed to an adequate portion of authentic language (Huei Lin, 2016). This view is supported by a myriad of meta-analyses arguing that by adopting DDL, learners are provided with higher access to accurate data, as corpora tend to give richer information compared with traditional resources (Boulton & Cobb, 2017; Cobb & Boulton, 2015; Leńko-Szymańska & Boulton, 2015).

Recent years have witnessed a growing academic interest in corpus linguistics. Barabadi and Khajavi (2017) posit that corpus is becoming a key instrument in teaching language. The reason why it is appealing is that it plays a vital role in language instruction (Boulton & Cobb, 2017; Chen et al., 2019; Leńko-Szymańska, 2017; Mizumoto & Chujo, 2016; Tono et al., 2014; Xie, et al., 2022). Therefore, corpora are broadly employed in language learning and instruction with diverse pedagogical aims and functions, such as preparing textbooks, dictionaries, vocabulary sources, and grammar books (Liou & Yang, 2020). The direct exploitation of corpora in language instruction referred to as DDL, has also attracted remarkable critical attention (Özer & Özbay, 2022; Vyatkina, 2016; Yılmaz, 2017; Yu, et al, 2022). For example, previous research has indicated DDL's potential to improve the teaching and learning of lexicogrammatical items (Huang, 2014). It further enhances learners' proofreading and error correction expertise (Mizumoto & Chujo, 2016). More broadly, it has been confirmed to improve cognitive and metacognitive skills (Yoon & Jo, 2014).

Many researchers verify the effectiveness of DDL for learning lexical bundles, vocabulary, collocations (Karras, 2016; Saeedakhtar et al., 2020; Tsai, 2019; Wu et al., 2019; Zhang, Akhter, and Al-Abyadh, 2022), and writing (Crosthwaite, 2020; Sun & Hu, 2020; Zhang, Akhter, Kumar, et al., 2022; Zhang, et al., 2022). Similarly, Mizumoto et al., (2017) and Sun and Hu (2020) suggest that students have positive reactions to the DDL approach. DDL approaches might also affect other cognitive and affective variables such as learners' learning autonomy and agency. Regarding learners' autonomy, DDL provides a potential means of enabling learners to gain control of their learning while instructors turn to facilitators (Tribble, 2012). Based on what was argued by Flowerdew (2015), DDL autonomy is reinforced by different groundbreaking theories such as constructivism (Brooks, 1993; Candy, 1991; Glaserfeld, 1995; Fox, 2001), noticing hypothesis of Krashen (1985), and sociocultural theory of Vygotsky (1978). As students examine the corpus, they become aware of particularities in linguistic patterns and then work cooperatively to recognize them. By collaborating, learners maximize chances to share their viewpoints and assist in meaningful negotiation with their peers. As a result, they are asked to give up their passive role and be involved in the process of meaning formation in collaboration with other participants.

On the other hand, with regard to learners' agency, it can be argued that it is first introduced by Bandura (2001). Gaining insight from social cognitive theory (SCT), it has been a much-debated issue in psychology and sociology for four decades. However, studies

highlighting the voice of learners' agency in English language acquisition are in their infancy. Agency, thus, refers to the capability of individuals to gain control over their cognition, motivation, and performance (Martin, 2004). Later on, Code (2010), gaining insight from SCT, proposes a model of the human agency known as the Agency for Learning (AFL) which encompasses an analysis of human agency through the lens of learners in educational environments. This model possesses four focal qualities of humans: intentionality, forethought, self-regulation, and self-efficacy. Intentionality implies a will to perform in a specific way according to mental state (Brady & Gilligan, 2019). Forethought entails the capacity to forecast the results of an action (Garcia-Martín & García-Sánchez, 2020). Self-regulation is described by Jansen et al. (2019) as a constructive procedure by which a learner determines a purpose for his/her learning and then tries to supervise and control his/her actions and cognitions directed by the set goals and the environmental factors. **Finally, self-efficacy** is defined as a functional self-correctness and self-awareness whereby learners think about their individual efficacy, beliefs, and behaviors to make necessary adjustments (Panadero et al., 2017).

DDL indeed has strong and rich theoretical and experimental background, but exploring the correlation between EFL Learners' perceived preferences for DDL and their learning agency, needs further investigation. This issue, which has not been appropriately investigated, is of much importance since learners' agency, directly and indirectly, affects several affective, cognitive, and academic variables. Based on this background, the following conceptual framework is provided:

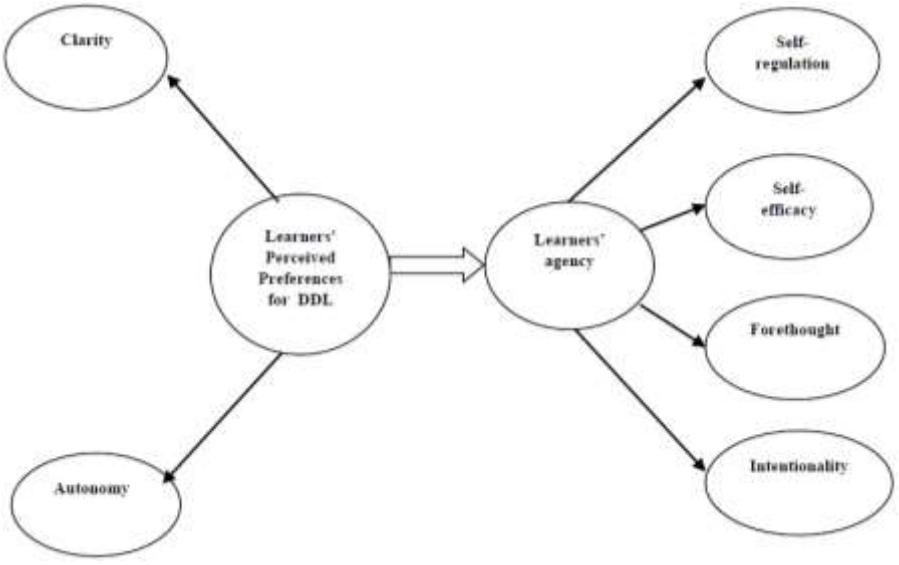

Figure 1. Structural and measurement model of Learners' Perceived Preferences for DDL and learners' agency

2. Purpose of the study

This study aims to explore the correlation between the Chinese EFL learners' perceived preferences for DDL and their learning agency. It also aims at exploring the correlation between learners' perceived preferences for DDL and different components of learners' agency. Therefore, this study aims at responding to the following stated questions.

1. Do the adopted scales of learners' agency, EFL learners' perceived preferences for DDL have acceptable levels of psychometrics?
2. Is there any significant correlation between EFL learners' agency and their perceived preferences for DDL?
3. Is there any significant correlation between components of EFL learners' agency and their perceived preferences for DDL?

3. Review of the Literature

Studies on Data-driven language learning

Review of the studies on DDL and its implications in foreign language classes have been studied by the researchers interested in English as a foreign language and Language for Specific Purposes (LSP) for two decades.

In theory, Johns (1991), the father of DDL, highlights the value of DDL and called "every student a Sherlock Holmes" (cited in Corino and Onesti, 2019, p.3). The main assumption behind this image as put by Corino and Onesti (2019) is that through searching linguistic data sources and online authentic language materials and sources, language learners have an access to grammatical patterns, lexical bundles, collocations, word meanings, and other aspects of language. Similarly, Boulton and Cob (2017) state that language corpora have definite implications in language classes, particularly for language learners and other users of English as EFL/ESL in an approach that is now called DDL. What's more, for assessing the benefits of DDL and learners' perceived preferences for DDL, Mizumoto (2016) develops and validate a 42-item scale which are reduced into four sub-scales: self-regulation, self-efficacy, forethought, and intentionality.

In practice, especially in the English writing class, using a quantitative case study, Corino and Onesti (2019) investigate the effect of

exposing LSP students to native speakers authentic oral and written language input. They find that students exposed to DDL did certain activities which considerably affect their use of language skills at work place. Furthermore, Liu (2016) has investigated the effects of DDL activities on EFL learners' writing development. He has reported that the experimental group outperformed the control group in writing accuracy and fluency. The qualitative results revealed the language learners' positive attitude toward using corpora but they had problems using it in the the writing process. By contrast, Yoon and Jo (2014), through a small case study, investigate their various positive effects on second language learners' error corrections in writing classes. They reveal that in indirect DDL, the rate of self-correction is higher than the correction rate in direct DDL for majority of the language learners, particularly for lower-level learners. They also stated that the direct DDL practices more positively influence the language learners' learner autonomy, particularly for learners with higher-level ability.This statement echoes the findings reported by Gaskell and Cobb (2004) indicating that novice or lower-level learners might find indirect DDL more appropriate.

On the basis, many reseraches explore the reason for the popularity of DDL.Previous studies (Boulton, 2009, 2011; Liu, 2016; Smart, 2014) on DDL verified that the most commonly cited aspects of DDL are learner-centered learning strategies and activities and authentic language data and corpus. Flowerdew (2015)has maintained that DDL has three main advantages. Firstly, it can provide language learners with authentic input. Secondly, DDL fosters learners' active engagement in the classroom activities and learning process, which promotes the language learners' ability to explore and discover language structures and rules based on their analysis, observation, and reflection on the concordance output. Finally, DDL fosters and promotes a more lexico-grammatical teaching approach by allowing the language learners to search for lexico-grammatical patters and structures in the corpus they are exposed to.

Studies on EFL Learners' Agency

In terms of the definition, Ahearn (2010) defines agency as "the socioculturally mediated capacity to act"(p.28). Human agency, as a theory, has been frequently used in humanities science such as sociology and psychology since the three decades ago (Ranjan et al., 2021). However, the research on EFL learners' agency is in developing phase (Maretha, & Waluyo, 2022). Ranjan et al. (2021) simply state that learners' agency can be used as a synonym for the actions learners do, which are deeply rooted in the learners' desires and motivations for learning and doing tasks. As mentioned by Allwright and Hanks (2009) agentic learners are commonly known as those active in controlling, monitoring, and constructing their learning.

Previous studies on the definition of learners' agency indicate that this research topic has been well investigated by several researchers because of education and instruction, curriculum, policy, and the students' societal needs (e.g., Chen, 2019; Fujieda, 2019; Ho, 2017; Jang, 2022). Therefore, a myriad of researches are carried out to explore the importance of learners' agency in English learning. Xiao (2014) suggest that agency can be important for learners to self-regulate their learning by assessing how an EFL learner brings their agency to play to foster language learning in terms of motivation, identity, motivation, self-efficacy and metacognition. Similarly, Liu and Chao (2017) reported how learners' agency was shown in classroom ecology. In another relevant study, Valdez et al. (2018) investigate the challenges and problems teachers in the Philippine education system face. They report that teachers believe that their learners are active agents and they could monitor and regulate their learning.

Learners' agency is also vital in the English writing class. In Korea, Jang (2022), through a qualitative study, explore Korean high school students' agency in the production of L2 processing and producing L2 writing. Field notes, online and offline interviews, and written materials over the course of two years are used as data collection instruments. The data are analyzed from a social view of agency. Results reveal that language learners' engagement in L2 writing projects varied, their awareness of linguistic resources is raised, and they reshape their L2 writing practices, develop their strategies placed resources, and their funds of knowledge. Similarly, Bhowmik (2016) has stated that agency is a decisive and important factor for L2 writers' reaction, performance, writing activities, and investment in writing processes. It has also been claimed that learner agency supports L2 writers' methods, strategies, and styles of writing (Cimasko & Shin, 2017; Jang, 2022; Jeffery & Wilcox, 2016).

In therms of the characteristics of learners' agency, Jang (2022) states that L2 learners' use of agentic factors is fluid rather than static, along which langue learner monitor and regulate the skills and strategies they employ to develop language skills, particularly writing skills. Very recently, Maretha, and Waluyo (2022) collect data from 389 undergraduate students. They analyze the data using descriptive statistics, and inferential statistics. Finally they find that the participants' learner agency vary across their language proficiency and their years of study. Learners' agency can be improved. Fujieda (2019) maintains that the participant recruited for the study have great academic progress by discursive processes, and through attending seminar research courses immersed into the discourse community and finally enhanced her agency.

4. Methodology

Participants

Participants of the study were selected among undergraduate English language learners majoring in English language literature, translation studies, and teaching English language. They were studying at universities in Shanghai, China. They were selected through convenience sampling. Both male (n=130) and female (n=100) undergraduate students were recruited. Their age ranged from 19 to 30 (M = 24.7, SD = 4). The students were aware of the researchers' intentions and objectives. They were filled in and signed the informed consent forms before the experiment, in which they were assured that their responses would be analyzed anonymously and would not negatively or positively affect their academic achievement

Instruments

For collecting the data, two scales were employed: one is long form of Agency for learning questionnaire (AFLQ) developed by Code (2020) and the other is learners' perceived preferences and benefits of Data-driven Learning developed by Mizumoto (2015). Learners' agency consists of 42 items, which are measured on a 5-point Likert scale (1= does not correspond to, 5= corresponds exactly). The questionnaire consists of four factors (intentionality=8 items, forethought=16 items, self-regulation=10 items, and self-efficacy =10 items). The second questionnaire consists of 16 items which were reduced to two factors (autonomy=8 items, clarity=8 items). The internal consistency of the scales was estimated using SPSS software, and Cronbach's alpha for the scales and their components exceeded 0.80, suggesting that the instruments enjoyed acceptable level of reliability (internal consistency).

Data Collection and Analysis Procedure

The study was taken in different steps. First, the participants were nominated and recruited. Second, the adopted questionnaires were administered to the recruited participants applications. The data were collected either through online applications or face-to-face meetings. Then, the questionnaires with missing data were excluded from the final analysis while the questionnaires with no missing data were coded and entered into SPSS. Four, the data were computed and transferred into AMOS software. The mean scores and standard deviations of the scores on the dependent and independent variables were calculated using SPSS. In addition, Analysis of Moment Structures (AMOS) software was used to test the proposed structural equation modeling (SEM) and the structural relations between the dependent variables (learner agency) and its components (intentionality, forethought, self-efficacy, and self-regulation) and the independent variable (learners' perceived preferences and benefits of DDL) and it its factors (clarity and autonomy). Finally, AMOS provides different model fit indices, among which the following indices were focused on the present research: CMIN (likelihood ratio chi-square test), Normed Fit Index (NFI), comparative fit index (CFI), adjusted goodness-of-fit index (AGFI), goodness-of-fit index (GFI), root mean square error of approximation (RMSEA) and X 2 /df.

5. Results

To test the hypothesis, multivariate analysis and structural equation modeling were conducted using Amos26 software. The results including preliminary descriptive statistics of the participants are presented first. Then, the results of inferential statistics for non-standard coefficients (significant coefficients) and standard coefficients (effect coefficients) of the measurement model and structural model of students' PPB of DDL and the learners' agency are presented in Figures 2 and 3 respectively.

Descriptive statistics

Results show that the participants' mean scores on the learners' agency variable, clarity, and autonomy are 3.86 (SD=1.1), 4.1 (SD=0.93), and 3.91 (SD=1.01) respectively, indicating that their learning agency level is acceptable. Results in Table 1 also show that the participants mean scores on the independent variable of learners' perceived preferences and benefits of DDL (M=3.6, SD=0.86), forethought (M=, SD=), self-efficacy (M=, SD=), self-regulation (M=, SD=), and intentionality (M=, SD) exceeded the hypothetical mean (M=2.5) suggesting the participants had positive preferences of DDL. (See Table 1)

Table 1. Mean and SD of participants' scores on research variables

Variables		Mean	SD
Learner agency	Clarity	3.86	1.12
	Autonomy	4.11	0.93
	Total	3.91	1.01
Learners' perceived preferences and benefits of DDL	Forethought	3.23	0.86
	Self-efficacy	3.91	0.83
	Self-regulation	3.84	0.91
	Intentionality	3.35	0.89
	Total	3.63	0.86

Results for Structural Equation Modeling (SEM)

Results of SEM analyzed through AMOS are presented in the following sections.

Inner Model Measurement Factor loadings of the dimensions of learners' agency and learners' perceived preferences and benefits of DDL based on non-standardized coefficients are presented in Table 2.

Table 2. Loading factors of the dimensions of the variables based on non-standard coefficients

Latent variable → observed variable	Loading factors (standard)	Standard Error	T	*p*
Autonomy →Learners PP & DDL	1.00			
Clarity→ Learners PP & DDL	1.29	0.22	5.78	0.001
Self-regulation → learner agency	1.00			
Self-efficacy → learner agency	0.668	0.05	12.93	0.001
Forethought → learner agency	0.91	0.08	10.48	0.001
Intentionality → learner agency	0.25	0.06	4.14	0.001

As shown in Table 2, all observed variables (dimensions of two independent and dependent variables) have a significant correlation

(P<0.01) with two independent and dependent variables as a total index (dependent variable) of students' perceived preferences and benefits of DDL and are active learners (that is, their factor loadings are meaningful).

Table 3. Correlation between the variables based on standard coefficients

Latent variable → observed variable	Loading factors (Standard coefficients)
Autonomy →Learners PP & DDL	0.66
Clarity→ Learners PP & DDL	0.95
Self-regulation → learner agency	0.91
Self-efficacy → learner agency	0.82
Forethought → learner agency	0.68
Intentionality → learner agency	0.30

As Figure 2 shows the measurement and structural model shows the effect of students' perceived preferences and benefits of DDL on their learner agency based on standard coefficients. Table 2, also shows that the amount of factor loadings (the amount of correlation between the dimensions of students' perceived preferences and benefits of DDL and learner agency) except for an intentionality dimension which showed a relatively weak correlation with the total indicators (the latent variables) are at a good level.

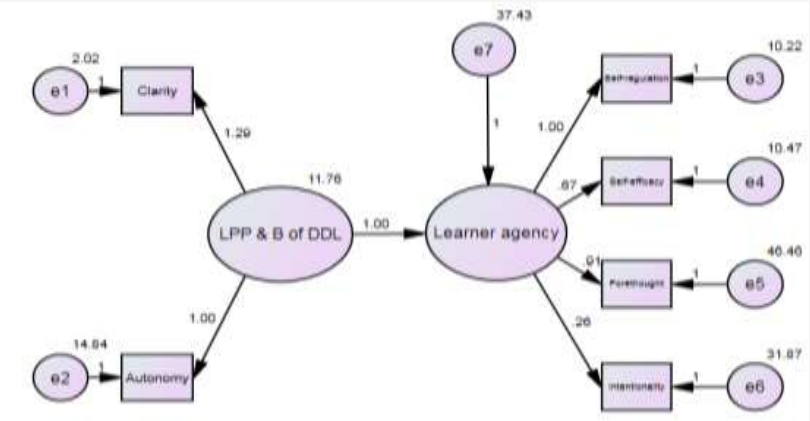

Figure 2. Structural model of LPP & B of DDL and learner agency (non-standard coefficients)

Note: LPP=learner perceived preferences and benefits, B of DDL= benefits of data-driven learning, **X^2 = 195/556, df =9, *p*= 0.001.**

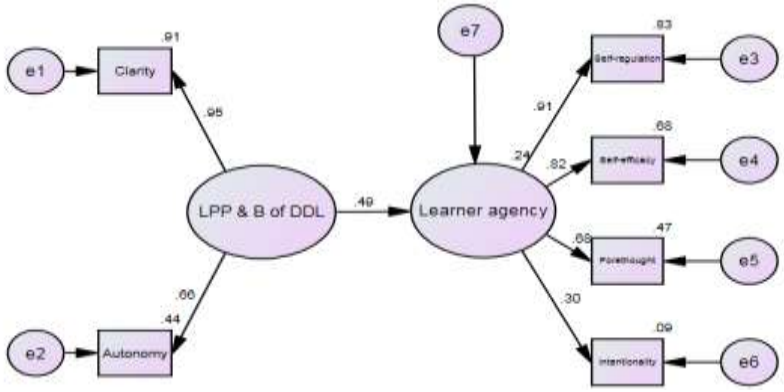

Figure 3. Structural model of LPP & B of DDL and learner agency (standard coefficients)

Structural model (path analysis)

Having performed the factor analysis, the researcher took the causal effects in the conceptual model of the research into consideration, to examine the effects of the observed variable on the latent variable. Table 3 shows the direct, indirect and total effects of the independent variable of students;' PPB of DDL on the dependent variable of learner agency in the form of standard coefficients:

Table 4. The direct and total effect of independent variable on learner agency

Independent variable	Learner agency		
	Direct effect	Indirect effect	Total effect
Learners' preferences and benefits of DDL	0.48	-----	0.48

As Table 4 shows, the students' PPB of DDL of DDL has a positive direct effect on the learners' learner agency. This indicates that for each

unit of increase in the independent variable 0.489 is added to the amount of the dependent variable of the learner agency.

General Evaluation of Model

The goodness of fit for the structural model of students' preferences and benefits of the DDL and their learner agency was estimated. Results are presented in Table 4.

Table 5. Results of Conformance of the Structural Model with Fitness Indicators

	CMIN	*p*	CMIN/DF	RMSEA	GFI	NFI	IFI	RFI	CFI	PCFI
Acceptable		>0.05	1-3	<0.05	>0.95	>0.95	>0.95	>0.90	>0.95	>0.60
Reported	1/55	0.001	5.00	0.065	0.93	0.91	0.930	0.365	0.92	0.57

As the findings in Table 5 show, the experimental structural and measurement model of the impact of students' preferences benefits of DDL on learner agency fits with the theoretical model. Therefore, according to the findings, it can be concluded that the research hypothesis indicating the developed theoretical model of structural relation between the students' PPB of DDL and the students' learner agency has acceptable goodness of fit.

6. Discussion

This study explored the structural relations between the language learners' PPB of DDL and the language learners' learner agency. The Chinese students attempted the questionnaires which measured the participants' perceived preferences of DDL and learner agency. AMOS software was used for analyzing the data. Results showed that the hypothetical model has acceptable goodness of Fit. That is, the independent variable and its components have significant effect the learners' agency. The loading factors in both standard and non-standard conditions verified the model fitness. It could be generally argued that learners' use of DDL affect their agency for learning. That is, exposure to language corpora affects their self-efficacy, intentionality, forethought, and self-regulation. The findings support the results of the previous related studies (Chang, 2014; Chambers and O'Sullivan, 2004; Gaskell and Cobb, 2004; Gilmore, 2009; Liu, 2016; Tono et al. 2014; Yoon, 2008) which concluded that DDL has a large number of benefits in teaching and learning a foreign language particularly writing skill. The findings also lend support to Chen and Flowerdew (2018, p. 355) argument suggesting that "DDL learners are given more responsibilities in exploring answers from the concordance by themselves, thus "mak[ing] the learning process appear to be more challenging and take longer."

Findings are also consistent with Corino and Onesti, (2019) who reported that exposing the learners of English for specific purposes to authentic linguistic corpora is beneficial and positively improves the ESP learners' use of authentic language in their actual working practice. Therefore, in line with the conclusions made by Flowerdew (2015), it could be strongly argued using DDL approaches in EFL/ESL classes has advantages and teachers are recommended to make the language learners aware of the significance of the DDL approaches and provide them with offline sources such as dictionaries and textbooks as well as online sources. Language learners are also suggested to use searching engines to find the needed corpora.

7. Conclusion

The study examined the measurement and structural model of EFL learners' preferences of DDL and their learning agency. The findings are particularly important and meaningful in that they ensure the researcher to verify that DDL approach is theoretically and pedagogically suitable and significant for enhancing the language learners' learner agency in general and their self-efficacy, self-regulation, and intentions for learning a foreign language in particular. However, the findings must be generalized with great care, as this study was undertaken in a unique context and the variables of the study were also unique. The other researchers are therefore recommended to replicate this study adding more dependent variables such as language skills and sub-skills.

References

Ahearn, L. M. (2010). *Agency and language.* In J. Jaspers, J. Östman, & J. Verschueren (Eds.), Society and language use (pp. 28-48). John Benjamin Publishing Company.

Allwright, D., & Hanks, J. (2009). *The developing language learner. An introduction to the exploratory practice*. Palgrave Macmillan. https://doi.org/10.1057/9780230233690

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26. https://doi.org/10.1146/annurev.psych.52.1.1

Barabadi, E., & Khajavi, Y. (2017). The effect of data-driven approach to teaching vocabulary on Iranian students' learning of English Vocabulary. *Cogent Education, 4*(1), 1283876. https://doi.org/10.1080/2331186X.2017.1283876

Bhowmik, S. K. (2016). Agency identity and ideology in L2 writing: Insights from the EAP classroom. *Writing & Pedagogy, 8*(2), 275-308. https://doi.org/10.1558/wap.26864

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning, 67*(2), 348-393. https://doi.org/10.1111/lang.12224

Brady, E., & Gilligan, R. (2019). The role of agency in shaping the educational journeys of care-experiences adults: insights from a life course study of education and care. *Child. Soc., 34*(2), 121-135. https://doi.org/10.1111/chso.12361

Charles, M. (2015). Same task, different corpus: The role of personal corpora in EAP classes. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 129-154). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.69.07cha

Chen, M., & Flowerdew, J. (2018). A critical review of research and practice in data-driven learning (DDL) in the academic writing classroom. *International Journal of Corpus Linguistics*, *23*(3), 335-369. https://doi.org/10.1075/ijcl.16130.che

Chen, M., Flowerdew, J., & Anthony, L. (2019). Introducing in-service English language teachers to data-driven learning for academic writing. *System, 87,* 102148. https://doi.org/10.1016/j.system.2019.102148

Chen, W. (2019). An exploratory study on the role of L2 collaborative writing on learners' subsequent individually composed texts. *Asian-Pacific Education Researcher, 28,* 563-573. https://doi.org/10.1007/s40299-019-00455-3

Cimasko, T., & Shin, D. S. (2017). Multimodal resemiotization and authorial agency in an L2 writing classroom. *Written Communication, 34*(4), 387-413. https://doi.org/10.1177/0741088317727246

Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 478-497). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139764377.027

Code, J. R. (2010). *Assessing agency for learning: Agency for learning and measuring agency for learning* [Unpublished Ph.D. dissertation]. Simon Fraser University.

Corino, E., & Onesti, C (2019). Data-Driven Learning: A Scaffolding Methodology for CLIL and LSP Teaching and Learning. *Front. Educ., 4*(7). https://doi.org/10.3389/feduc.2019.00007

Flowerdew, L. (2015). Data-driven learning and language learning theories: Whither the twain shall meet. In A. Leńko- Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 15-36). John Benjamins Publishing Company. https://doi.org/10.1075/scl.69.02flo

Fujieda, Y. (2019). Academic discourse socialization in a research seminar course: A case study of a Japanese EFL undergraduate learner. *Asian-Pacific Education Researcher, 28,* 93-100. https://doi.org/10.1007/s40299-018-0416-z

Garcia-Mart ń, J., & Garc á-S ánchez, J. N. (2020). The effectiveness of four instructional approaches used in a MOOC promoting personal skills for success in life. *Revista de. Psicodid áctica, 25*(1), 36-44. https://doi.org/10.1016/j.psicoe.2019.08.001

Ho, M. C. (2017). Navigating scholarly writing and international publishing: Individual agency of Taiwanese EAL doctoral students. *Journal of English for Academic Purposes, 27,* 1-13. https://doi.org/10.1016/j.jeap.2017.02.004

Huang, J. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL*, *26*(2), 163-183. https://doi.org/10.1017/S0958344014000020

Huei Lin, M. (2016). Effects of corpus-aided language learning in the EFL grammar classroom: A case study of students' learning attitudes and teachers' perceptions in Taiwan. *TESOL Quarterly, 50*(4), 871-893. https://doi.org/10.1002/tesq.250

Jang, J. (2022). An exploratory study on learner agency and second language writing practices of Korean high school students. Asian. J. Second. *Foreign. Lang. Educ. 7,* 31. https://doi.org/10.1186/s40862-022-00158-1

Jansen, R., van Leeuwen, A., Janssen, J., Jak, S., & Kester, L. (2019). Self-regulated learning partially mediates the effect of self-regulated learning interventions on achievement in higher education: a meta-analysis. *Educational Research Rev*iew, *28,* 100292. https://doi.org/10.1016/j.edurev.2019.100292

Jeffery, J. V., & Wilcox, K. C. (2016). L1 and L2 adolescents' perspectives on writing within and across academic disciplines: Examining the role of agency in writing development. *Writing & Pedagogy, 8*(2), 245-274. https://doi.org/10.1558/wap.28750

Johns, T. (1991). Should you be persuaded: two samples of data-driven learning materials. *English Language Research Journal*, *4,* 1-16.

Jones, C., & Waller, D. (2015). *Corpus linguistics for grammar a guide for research*. Routledge. https://doi.org/10.4324/9781315713779

Leńko-Szymańska, A. (2017). Training teachers in data-driven learning: Tackling the challenge. *Language, Learning and Technology, 21*(3), 217-241.

Leńko-Szymańska, A., & Boulton, A. (2015). Multiple affordances of language corpora for data-driven learning. *International Journal of Corpus Linguistics*, *20*(4), 560-569. https://doi.org/10.1075/ijcl.20.4.07gar

Lin, M. H., & Lee, J. Y. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *ELT Journal*, *69*(3), 264-274. https://doi.org/10.1093/elt/ccv010

Liou, H., & Yang, T. (2020). Data-driven learning at the English drafting stage. In M. Kruk & M. Peterson (Eds.), *New technological applications for foreign and second language learning and teaching* (pp. 283-305). IGI Global. https://doi.org/10.4018/978-1-7998-2591-3.ch014

Liu, Q., & Chao, C. C. (2017). CALL from an ecological perspective: How a teacher perceives affordance and fosters learner agency in a

technology-mediated language classroom. *ReCALL, 30,* 1-20. https://doi.org/10.1017/S0958344017000222

Maretha, A. L., & Waluyo, B. (2022). Profiles, differences, and roles of learners' agency in english learning in Thailand. *International Journal of Instruction, 15*(2), 543-564. https://doi.org/10.29333/iji.2022.15230a

Martin, J. (2004). Self-regulated learning, social cognitive theory, and agency. *Educ. Psychol., 39,* 135-145. https://doi.org/10.1207/s15326985ep3902_4

Mizumoto, A., & Chujo, K. (2016). Who is data-driven learning for? Challenging the monolithic view of its relationship with learning styles. *System, 61,* 55-64. https://doi.org/10.1016/j.system.2016.07.010

Mizumoto, A., Chujo, K., & Yokota, K. (2016). Development of a scale to measure learners' perceived preferences and benefits of data-driven learning. *ReCALL, 28*(2), 227-246. https://doi.org/10.1017/S0958344015000208

Özer, M., & Özbay, A. S. (2022). Exploring the data-driven approach to grammar instruction in the ELT context of Turkey. *Australian Journal of Applied Linguistics, 5*(2), 35-63. https://doi.org/10.29140/ajal.v5n2.713

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: four meta-analyses. *Educational Research Review, 22,* 74-98. https://doi.org/10.1016/j.edurev.2017.08.004

Poole, R. (2016). A corpus-aided approach for the teaching and learning of rhetoric in an undergraduate composition course for L2 writers. *Journal of English for Academic Purposes, 21*(2), 99-109. https://doi.org/10.1016/j.jeap.2015.12.003

Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL, 26,* 184-201. https://doi.org/10.1017/S0958344014000081

Tono, Y., Satake, Y., & Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL, 26*(2), 147-162. https://doi.org/10.1017/S095834401400007X

Tribble, C. (2012). Corpora in the language teaching classroom. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1377-1379). Blackwell. https://doi.org/10.1002/9781405198431.wbeal0226

Tribble, C. (2015). Teaching and language corpora: Perspectives from a personal journey. In A. Lenko-Szymanska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 37-62). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.69.03tri

Valdez, P. N., Navera, J. A., & Esteron, J. J. (2018). What is reflective teaching? Lessons learned from ELT teachers from the Philippines. *Asian-Pacific Education Researcher, 27,* 91-98. https://doi.org/10.1007/s40299-018-0368-3

Vyatkina, N. (2016). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology, 20*(3), 159-179.

Xiao, J. (2014). Learner agency in language learning: The story of a distance learner of EFL in China. *Distance Education, 35*(1), 4-17. https://doi.org/10.1080/01587919.2014.891429

Xie, D., Saeed, N., Akhter, S., & Kumar, T. (2022). A step towards a sustainable environment in top Asian countries: the role of higher education and technology innovation. *Economic Research-Ekonomska Istraživanja,* 1-24. https://doi.org/10.1080/1331677X.2022.2152359

Yılmaz, M. (2017). The effect of data-driven learning on EFL students' acquisition of lexico-grammatical patterns in EFL writing. *EJAL Eurasian Journal of Applied Linguistics, 3*(2), 75-88. https://doi.org/10.32601/ejal.460966

Yoon, H., & Jo, J. W. (2014). Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing. *Language Learning & Technology, 18*(1), 96-117.

Yu, G., Akhter, S., Kumar, T., Ortiz, G. G. R., & Saddhono, K. (2022). Innovative application of new media in visual communication design and resistance to innovation. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.940899

Zhang, D., Akhter, S., Kumar, T., & Nguyen, N. T. (2022). Lack of Emotional Experience, Resistance to Innovation, and Dissatisfied Musicians Influence on Music Unattractive Education. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.922400

Zhang, X., Akhter, S., Nassani, A. A., & Haffar, M. (2022). Impact of News Overload on Social Media News Curation: Mediating Role of News Avoidance. *Frontiers in Psychology, 13,* 865246-865246. https://doi.org/10.3389/fpsyg.2022.865246

Zhang, Z., Akhter, S., & Al-Abyadh, M. A. (2022). Determinants of Unverified News Sharing on Social Media and Its Effects on Corporate Image. *Frontiers in Psychology,* 3583. https://doi.org/10.3389/fpsyg.2022.937104