ISSN 2411-7390

jle

JOURNAL OF LANGUAGE & EDUCATION

Volume 8 Issue 3, 2022

HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

EDITOR-IN-CHIEF

EDITORIAL BOARD

EDITORIAL TEAM

CONTENT

EDITORIAL

RESEARCH ARTICLES

PILOT STUDY

REVIEW PAPER

REVIEWS

https://doi.org/10.17323/jle.2022.15960

Citations and References: Guidelines on Literature Practices

Elena Tikhonova[1, 2], **Lilia Raitskaya**[3]

[1] National Research University Higher School of Economics, Moscow, Russia

[2] Peoples' Friendship University of Russia (RUDN University), Moscow, Russia

[3] Moscow State Institute of International Relations (MGIMO University), Moscow, Russia

ABSTRACT

Introduction. Citations of scholarly publications are considered an efficient measure of productivity of research and researchers. They are part of scholarly communication, driving the evolving knowledge in all disciplines. Citations form an integral part of literature practices of researchers. The latter are prone to deliberate or unconscious biases. One of the challenges all researchers face is to overcome or at least mitigate identified biases in citation. It may lead to distorting knowledge development in the least possible way.

Purpose. As the research of citations is on the rise, more theoretical background is being developed. The editors call the JLE readers attention to the Triangular Citations, with various relationships among publications formed by citations.

Results. Empirical studies are not unanimous in defining the features that affect citations most. Most researchers highlight a publication source as the most influential feature, with multi feature approach favoured by the academic community at large. Other features entail features relating to authors, journal, and publication itself. Moreover, there are varieties across the disciplines. The field of research of citation behaviour is only beginning to widen. The focus of such research is made on citation behaviour patterns, individual incentives behind the references and citations aligned with financial or reputation stimuli, and citations patterns linked to citation behaviour. Bias in citation threatens to distort knowledge and may evolve it subtly or obviously in a specific direction. Biases are not easy to deal with. In addition, values and mindsets vary across the countries and academic and scholarly communities that hinders efforts to overcome biases.

Implications. The JLE editors sum up the best guidelines on improvements in publications that add to greater citations, with the high quality of articles as the key.

KEYWORDS:

citation, citation behaviour, citation bias, citation recommendation, literature practice, citation network, triangular citation, citation stacking, citation boosting.

INTRODUCTION

Scholarly publishing is inseparable from citation issues. The latter are getting more influence and weight for universities in respect of university rankings (Waltman et al., 2012); for scholars and academics at large as citations and H-index affect their academic standings and contracts of employment; and for scholarly journals as a key indicator of their prominence within their scopes. Assessments of research and researchers in the academia rank citations high on the list (Olszewski, 2020). Citation content of research papers is "an important representation of communication among different disciplines" in social sciences and humanities that are inevitably multidisciplinary fields (Zhang, Liu, & Wang, 2021).

A competitive edge in science adds much to new jobs and economic growth. Science bringing innovations into the economy acts as a production factor. Thus, funding research came to the fore some twenty plus years ago in most economies. It is essential for funders to be sure that their investments would pay off (Bu et al., 2021). Being nearly uniform so far, global trends "publish or perish" embody policies of the countries seeking innovations and cutting-edge research. Formally, such research publications are often highly cited.

Citation: Tikhonova, E., & Raitskaya, L. (2022). Citations and References: Guidelines on Literature Practices. *Journal of Language and Education, 8*(3), 5-10. https://doi.org/10.17323/jle.2022.15960

Correspondence:
Elena Tikhonova,
etihonova@hse.ru

Received: June 28, 2021
Accepted: September 10, 2022
Published: September 30, 2022

Citations of the published research results have become the major and most reliable measure of research impact and consequently their value. As peer-reviewed journals are the prevailing form of scholarly communication where new research results are published after peer review, they get accessible for the global scholarly communication there. But so far approximately 90 percent of the studies published globally are never cited (Meho, 2007). Even articles that are read, downloaded, or publicly mentioned in the social media remain not mentioned by the scholarly community. To make their publications citable, authors ought to know what and why is cited.

As science is supposed to be communal in nature, no research is conducted on its own as a stand-alone study. All new research publications are based on research of other scholars. The previously published and cited, i.e. acknowledged, research papers lay the foundation for new research. Science funders are eager to have the publications they financially support brought out in highly reputed journals and consequently cited. It became the objective behind the "publish or perish" policies.

In their studies, researchers spend much time on literature practices. The latter ranges from searches for scholarly information of different kinds to analyses and citation of articles and monographs on the topic of interest. Literature practices "consists of the selection, reading and citing of sources" (Klitzing, Hoekstra, & Strijbos, 2019).

Many biases in citations stem from inefficient, prejudiced, or wrong literature practices. Citation behaviour patterns such as self-citation boosting and citation stacking also may distort the ways of citing a source (de Lusignan & Moen, 2016).

Citations and refences gave birth to a field covering various notions and terms (see the box below). For successful authors, it is challenging to realize what their papers should be to be both socially optimal and highly citable. The field is supplying them with some clues and recommendations.

The main part of this editorial entails an analysis of publication features driving citation growth, prevailing citation behaviour patterns and motivations lying behind; missteps and bias in literature practices; and a theoretical basis of citation networks. Finally, we outline essential guidelines for JLE authors on avoiding bias in citations and their completeness, using the best literature practices and key characteristics of the existing citation recommendation systems.

Publication Features Affecting Citations

Researchers of citation single out various sets of features driving citation growth. All researchers are unanimous in multi-feature approach. The major factors entail inherent quality of a scholarly publication; features relating to authors, including the number of authors, their academic reputation, and nationality or affiliation; the impact of the journals where an article is published (journal centrality and SCImago quartile, impact factor, language, etc.); paper features (document type, topic, its rank, title, number and prestige/ reliability of references; novelty, etc.) (Bai, Zhang, & Lee, 2019). This list is too general and may be thoroughly extended.

In some studies, authors outline measurable citations, with a publication source as the most important factor contributing to citations (Ha, 2022). Only measurable variables were included in the model in the research by T. Ha (2022). They were analysed to estimate the ways they affect the number of citations. The variables entail the publication year, number of authors of the paper, number of author affiliations, number of author countries, number of pages, number of words in the title, number of words in the abstract, number of author keywords, number of references, number of All Science Journal Classification (ASJC) codes, publication source, form of open access, publication type, and language (Ha, 2022).

Recurrent authors tend to provide for special features in their publications in pursuit of citations. Our personal experience involves novelty in the mainstream topics (via a unique approach or novel aspects of a popular topic); types of papers (reviews are generally cited more often as compared with other types); the reputation of a journal; and ingenious methods or design of research. There is a reservation for any set of prevailing features. It is the subject field of research. The features specific of a particular discipline may be rather distinctive and differ a lot from the stereotyped characteristics across other disciplines (Vanday, 2009).

Citation Behaviour Patterns

Citation behaviour describes the ways an author selects references and disseminate knowledge. It is thoroughly in-

• citation	• citer	• citation stacking
• quotation	• citation pattern	• citation boosting
• quote	• self-citation	• bibliographic coupling
• refence	• citation network	• citation frequency
• paraphrasing	• co-citation	• citation index
• literature practice	• direct citation	• h-index
• triangular citation	• in-text citation	• citation-seeking behaviour
• topic prominence	• accuracy of citation	• citation-driven research
• citation recommendation system	• redundancy citations	• notification of new citations
• replication crisis	• secondary citation	• shortcut
	• impact factor	• refence linking, etc.

vestigated in two approaches: authors and citation contexts (Yang, & Liu, 2022). Relationships among research papers are constructed, with citation networks and citation cascades prevailing elsewhere.

Citation behaviour is supposed to be the source driver of scientific dynamics (Yang, & Liu, 2022). On the whole, authors may choose to follow one or several behaviour patterns in citing scholarly literature. Citation behaviour is defined as "the author's selection of references based on the motivation of knowledge expression" (Zhang, Ding, & Milojevic, 2013). But this definition is limited to a great extent to the normative theories where citations are considered based merely on topical and intellectual grounds. It is obvious that other motivations and incentives may come into play.

The motivations of authors seeking frequent citations are diverse. Individual incentives are essentially aligned with reputation and financial issues. Those researchers who shop for citations may easily make missteps while searching for literature or referencing to what would be advantageous. In addition, they tend to sacrifice social value, preferring research that would be cited more frequently (Olszewski, 2020).

To improve the quality of citations, researchers should increase their awareness of citation patterns. In linguistics, citation patterns are studied on their own by their rhetorical contexts. The analysis covers "densities, surface forms, roles of cited authors, reporting verbs, and functions" (Zhang, 2022).

Bias in Literature Practices

Avoiding bias in science is one of the strategic aims. Bias at any stage of research threatens to distort knowledge and impair the objectivity of research. Biases emerge at various stages of the research process, including citations. But bias is not always a measurable concept. Moreover, biases in citations may be influenced by prejudices in other aspects of research. Citing publications selectively may drive knowledge development "subtly into a certain direction" (De Vries, 2018).

What constitutes a citation bias? What skewed citation distributions prevail in research? As citations serve as a confirmation of contribution to science by the researchers' community, a citation bias is considered as a confirmation bias. Such a bias "connotes the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand" (Nickerson, 1998: 175). Citation distributions may be skewed in various aspects (language, gender, nationality, schools of thought, journal-related citations, etc.). Though, one bias persists in all fields of knowledge. Most researchers prefer to cite sources in their native language (Bookstein, & Yitzhaki, 1999). This bias is more evident and far from being subtle as other biases tend to be.

There is also evidence that many authors select and cite sources and documents, not judging by their content. They use references to prove their ideas by citing reputed researchers in their field (Tahamtan, & Bornmann, 2018). This motivation may add to a bias.

Some time ago, gender imbalances in neuroscience reference lists made several researchers study the situation (Dworkin et al., 2020). They found that there was a gender bias as the imbalance in reference lists was "statistically unexpected" whereas "gender was not a factor" (Zurn, Bassett, & Rust, 2020). To mitigate the bias, the authors put up a special statement to be incorporated in research texts. The Citation Diversity Statement is aimed at increasing awareness of possible citation bias (Zurn, Bassett, & Rust, 2020). This practice is not common so far and not easy to follow.

Values and mindsets are not uniform and far from similar across the world. They cannot be forced by any researchers' communities upon other groups of researchers as science is undoubtedly a domain of civil freedoms. In some parts of the world, bias is approached differently, with a citation diversity statement put forward to mitigate alleged biases in citations of some social or other groups of researchers. Though, in other countries the only criterion of bias is related to the ignored quality of research and scope of the field. In those countries, personalities and social features of authors cited elsewhere are considered regardless of science and citations.

Some Theory on Citation Networks: Triangular Citations

Following the science citation indices introduced in the 1960s, the citation analysis field was beginning to evolve. The latter entails analyses of the key citation relationships – direct citations, co-citation, and bibliographic coupling (Liu, Yang, & Chen, 2021). Such relationships on bigger databases lead to citation networks and citation cascades.

The more citations several publications share, the closer and the more similar the underlying studies are. So called "triangular citations" are based on triangular of references, also known as "triangular of meaning". Originally, it is a model explaining how linguistic symbols relate to the objects they represent. The model of 1923 by C.K. Odgen and I.A. Richards. It essentially reminds of the concepts previously offered by B. Bolzano (1810) and Aristotle (4th century B.C.). The relationships in the citation-applied model are shown in Diagrams 1 and 2.

Citation networks describe the relationships among some group of the documents or a database. The triangular model explains the levels of relationships and their interaction. Citation networks, or citation graphs are widely used in citation analysis. Whereas citation cascade is "the constitution

Diagram 1

Citation Triangular

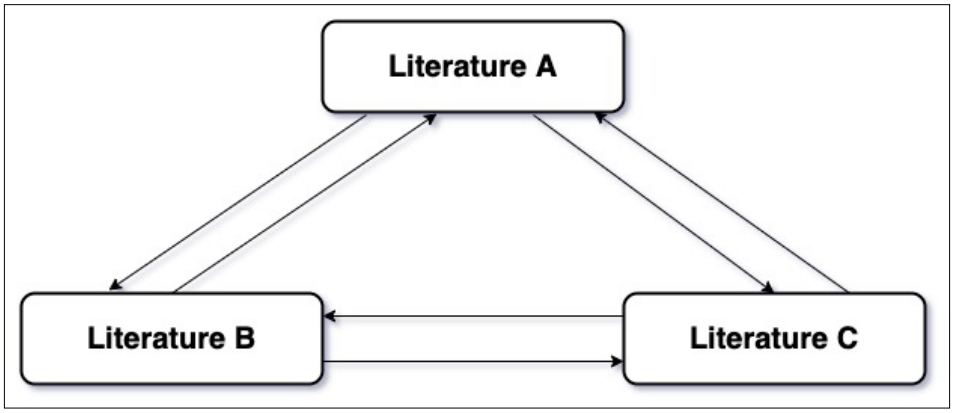
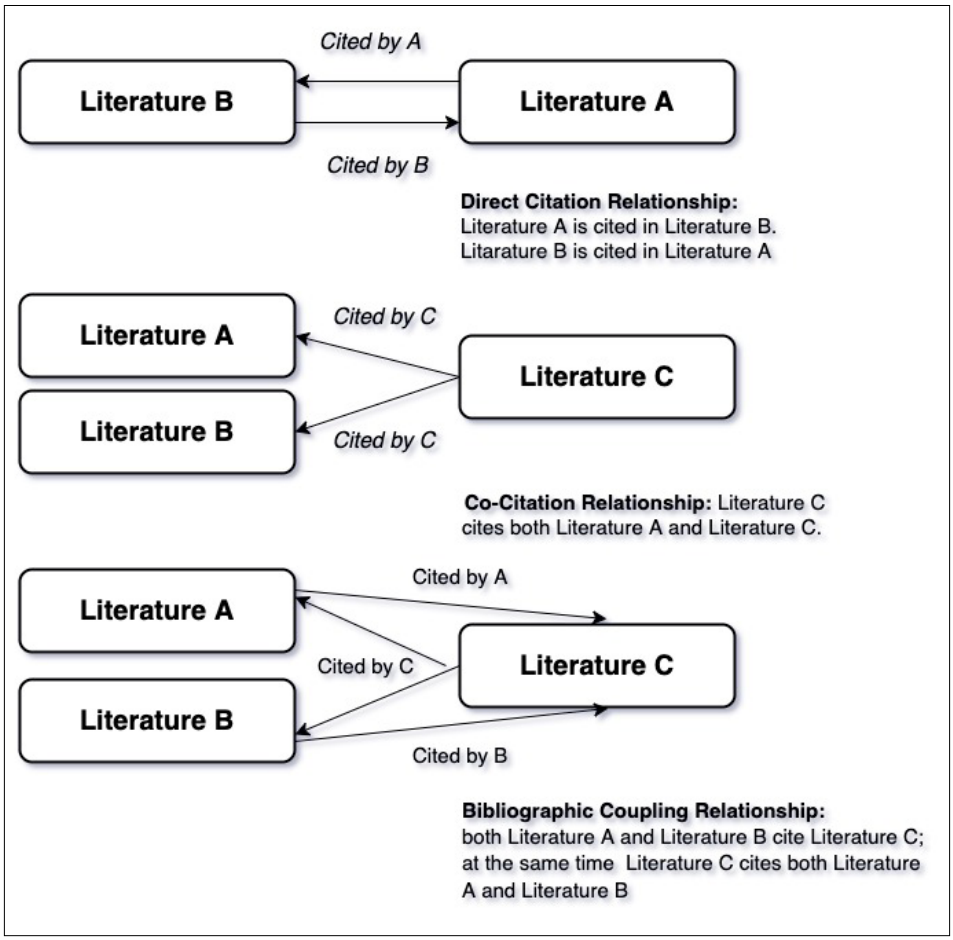


Diagram 2

Citation Triangular: Citation Relationships



of a series of subsequent citing events initiated by a certain publication" (Min et al., 2020).

Researchers of citations outline the core elements in the process of citing publications. Based on empirical studies, they include the context of the cited document, processes from selection to citation of documents, and the context of the citing document (Tahamtan, & Bornmann, 2018)

Conclusions and Applications for JLE Authors

The JLE as a scholarly peer-reviewed journal is making its utmost to overcome or avoid any biases in citations and any other aspects of scholarly communication. The JLE keeps their recurrent and future authors abreast of any practices that may improve their submissions to the JLE. Based on the above, we would invite our authors to focus more of their efforts on the objective and unbiased selection of sources to

cite. Though it may be a challenge to overcome a bias produced by the language repertoire of a researcher, the cited literature in the reference lists must represent the scope of the field to its advantage.

We realise that some issues are studied locally or regionally. If it is the case, all efforts ought to be made to find out whether there are any international or foreign publications associated this such local or reginal cases. Our reviewers tend to take a responsible attitude to the reference lists of all submissions to the JLE. Part of submissions are desk-rejected merely on the ground of incomplete or insufficient lists. It always implies a serious bias in citations.

The JLE expects that its future submissions would include citations of not only the prominent studies previously published and acknowledged by the scholarly community, but the cutting-edge research containing brand-new ideas, even still disputable. The JLE as any journal is aimed at producing a heated scientific discussion of new concepts.

In conclusion, the JLE welcome submissions on the raised issues, especially the linguistic patterns of citations and bias in citations. Further research may add a lot to fill the gaps that still exist in this field.

DECLARATION OF COMPETING INTEREST

None declared.

REFERENCES

Bai, X., Zhang, F., & Lee, I. (2019). Predicting the citations of scholarly paper. *Journal of Informetrics, 13*(1), 407-418. http://doi.org/10.1016/j.joi.2019.01.010

Bookstein, A., & Yitzhaki, M. (1999). Own-language preference: A new measure of "relative language self-citation". ***S**cientometrics, 46*(2), 337-348. http://doi.org/10.1007/BF02464782

Bu, Y., Waltman, L., & Huang, Y. (2021). A multidimensional framework for characterizing the citation impact of scientific publications. *Quantitative Science Studies, 2(*1), 155-183. http://doi.org/10.1162/qss_a_00109

Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. ***N**ature Neuroscience, 23(*8), 918-926. https://doi.org/10.1038/ s41593-020-0658-y

Ha, T. (2022). An explainable artificial-intelligence-based approach to investigating factors that influence the citation of papers. *Technological Forecasting and Social Change, 184* http://doi.org/10.1016/j.techfore.2022.121974

Klitzing, N., Hoekstra, R., & Strijbos, J. (2019). Literature practices: Processes leading up to a citation. *Journal of Documentation, 75*(1), 62-77. http://doi.org/10.1108/JD-03-2018-0047

Liu, Y., Yang, L., & Chen, M. (2021). A new citation concept: Triangular citation in the literature. *Journal of Informetrics, 15*(2). http://doi.org/10.1016/j.joi.2021.101141

de Lusignan, S., & Moen, A. (2016). Extracting oneself from the citation-stacking bear trap. *Methods of Information in Medicine, 55*(4), 301-302. http://doi.org/10.3414/ME16-03-0001

Meho, L. I. (2007). The rise and rise of citation analysis. *Physics World, 20*(1), 32-36. http://doi.org/10.1088/2058-7058/20/1/33

Min, C., Chen, Q., Yan, E., Bu, Y., & Sun, J. (2020). Citation cascade and the evolution of topic relevance. *Journal of the Association for Information Science and Technology.* http://doi.org/10.1002/asi.24370

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175-220. http://doi.org/10.1037/1089-2680.2.2.175

Olszewski, W. (2020). A theory of citations. *Research in Economics, 74*(3), 193-21. http://doi.org/10.1016/j.rie.2020.06.001

Tahamtan, I., & Bornmann, L. (2018). Core elements in the process of citing publications: Conceptual overview of the literature. *Journal of Informetrics, 12*(1), 203-216. http://doi.org/10.1016/j.joi.2018.01.002

Vanclay, J. K. (2009). Bias in the journal impact factor. *Scientometrics, 78*(1), 3-12. http://doi.org/10.1007/s11192-008-1778-4

De Vries, Y. A., Roest, A. M., De Jonge, P., Cuijpers, P., Munafò, M. R., & Bastiaansen, J. A. (2018). The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: The case of depression. *Psychological Medicine, 48*(15), 2453-2455. http://doi.org/10.1017/S0033291718001873

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., Van Eck, N. J., Van Leeuwen, T.N., Van Raan, A.E., & Wouters, P. (2012). The Leiden ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology, 63*(12), 2419-2432. http://doi.org/10.1002/asi.22708

Yang, J., & Liu, Z. (2022). The effect of citation behaviour on knowledge diffusion and intellectual structure. *Journal of Informetrics, 16*(1). http://doi.org/10.1016/j.joi.2021.101225

Zhang, G. (2022). The citational practice of social science research articles: An analysis by part-genres. *Journal of English for Academic Purposes, 55*. http://doi.org/10.1016/j.jeap.2021.101076

Zhang, G., Ding, Y., & Milojevic, S. (2013). Citation content analysis (CCA): A framework for syntactic and semantic analysis of citation content. *Journal of the American Society for Information Science and Technology, 64*(7), 1490-1503. http://doi.org/10.1002/asi.22850

Zhang, C., Liu, L., & Wang, Y. (2021). Characterizing references from different disciplines: A perspective of citation content analysis. *Journal of Informetrics, 15*(2). http://doi.org/10.1016/j.joi.2021.101134

Zurn, P., Bassett, D. S., & Rust, N. C. (2020). The citation diversity statement: A practice of transparency, A way of life. *Trends in Cognitive Sciences, 24*(9), 669-672. http://doi.org/10.1016/j.tics.2020.06.009

https://doi.org/10.17323/jle.2022.12750

Rhetorical Structure of Applied Linguistics Research Article Discussions: A Comparative Cross-Cultural Analysis

Leila Ahmadi

Volgograd State University, Volgograd, Russia

ABSTRACT

Background. Recent years have seen tremendous research efforts in the development of English for academic and research publication purposes, utilising an established approach to comparative genre analysis. This growing interest is primarily driven by the global dominance of Anglophone writing conventions, which necessitates raising awareness among researchers, particularly in non-Anglophone contexts.

Purpose. This study explored and analysed the communicative intents of the discussion sections in research articles in two different contexts to investigate the effect of nativeness on the structural organisation in this genre. The focus of the study was on the rhetorical structure and employment of Moves in the applied linguistics research article Discussions, written in English by Iranian and native English-speaking researchers.

Methods. A mixed-methods research study was conducted on two corpora, comprising 40 Discussions written by Iranian scholars and 40 Discussions written by native English-speaking scholars, selected from research articles published in international peer-reviewed journals.

Results. The comparison of the two corpora revealed similarities and differences in the frequency, type, structure, sequence, and cyclicity of Moves. While there were significant differences in the frequency and sequence of Moves and Steps, both corpora employed the same types. They featured cyclical structures with no evidence of linear patterns across the Discussions. Both groups of researchers found it essential to provide background information and report and comment on the results in the research article Discussions, however, with notable differences in commenting strategies, i.e., Steps. The results indicated that socio-cultural conventions might have influenced the scholars' under- and over-employment of certain Moves and Steps in the research article Discussions.

Implications. The findings of this study provide research-based evidence to practically and pedagogically assist in the context of English for academic and specific purposes, particularly in teaching English for research publication purposes to non-native English-speaking scholars.

KEYWORDS:

research article, discussion section, academic writing, communicative purpose, genre-based study, move analysis, rhetorical organisation, applied linguistics

Citation: Ahmadi, L. (2022). Rhetorical Structure of Applied Linguistics Research Article Discussions: A Comparative Cross-Cultural Analysis. *Journal of Language and Education, 8*(3), 11-22. https://doi.org/10.17323/jle.2022.12750

Correspondence:
Leila Ahmadi,
leilaahmadi@outlook.com

Received: July 18, 2021
Accepted: September 17, 2022
Published: September 30, 2022

INTRODUCTION

Research publication facilitates spreading, preserving, and transmitting knowledge to the next generation. Genre analysis has, therefore, paid significant attention to research articles (RAs) over the past several decades. Successful academic writing is as much the result of appropriate rhetorical and linguistic choices as the quality of the content. Hence, much research has been conducted to ensure quality research publications. English is considered the principal medium of scientific communication and publication. Consequently, academic writing is dominated by Anglophone conventions in the global context, and non-native scholars are increasingly willing to communicate their ideas in English to gain international recognition and enhance their scholarly impact. It has been widely recognised that writing for scholarly publication in English is a critical dimension of national and international academic and professional standing

(Stoller & Robinson, 2013; Yang, 2016). However, complying with the requirements of the genre in RAs is significantly more challenging for non-native English speakers (hereafter NNES) than native English speakers (hereafter NES) (Burrough-Boenisch, 2003; Cho, 2004; Sadeghi & Shirzad Khajepasha, 2015), even for expert professionals (Glasman-deal, 2020; Lei & Hu, 2019). The underlying reason for this issue is partly due to the undeniable impact of cultural and contextual factors on the form and structure of academic writing. In addition, insufficient knowledge of rhetorical structures and the Anglophone conventions concerning the semantic, syntactic, and pragmatic choices in the presentation of information places NNES at a significant disadvantage against NES for the publication of RAs in English.

To express their communicative intent, researchers use a variety of structural organisations for which the information is segmented into functional units of discourse, commonly referred to as "Move". Moves are units of discourse structure that have a uniform orientation, a particular structure, and clearly defined functional characteristics (Nwogu, 1989). It has been a significant focus of recent scholarship to examine and dissect the rhetorical structure of the canonical sections of scientific papers. Several studies have explored and presented frameworks for all sections of RAs (e.g., Kanoksilapatham, 2005; Lewin et al., 2001; Nwogu, 1989), while others have focused on specific sections such as the introduction (e.g., Samraj, 2002; Swales, 1990), methods (e.g., Chen & Kuo, 2012; Lim, 2006), results (e.g., Brett, 1994), and discussion (e.g., Basturkmen, 2012; Peacock, 2002). As a text analytical approach, many researchers have adopted Move analysis to identify the structure of information in different contexts, registers, and genres, such as academic prose (e.g., Alamri, 2020; Alinasab et al., 2021; Hu & Liu, 2018; Loi et al., 2016; Marefat & Mohammadzadeh, 2013). Researchers classified Moves as conventional if they appeared in 50% (Holmes, 1997; Swales, 1990), 60% (Kanoksilapatham, 2005), or 66% (Hatzitheodorou, 2014), and obligatory if they appeared in 90% (Santos, 1996) or 100% (Cross & Oppenheim, 2006; Kanoksilapatham, 2005) of the investigated genre.

Given that the importance and contribution of findings are established in the discussion section of RAs (Ruiying & Allison, 2003), extensive research has employed Move analysis to investigate the rhetorical structure of this section (e.g., Amirian et al., 2008; Amnuai, 2019; Arabi, 2019; Atai & Falah, 2005; Basturkmen, 2012; Peacock, 2002). Hopkins and Dudley-Evans (1988) conducted an early study on this subject. They proposed an eleven-Move scheme based on their detailed analysis of the discussion section of theses and published articles in biology. In his Create a Research Space (CARS) model, Swales (1990, 2004) glossed over the eleven-Move schemes suggested by Peng (1987) and Hopkins and Dudley-Evans (1988), proposing a list of the eight most frequent Moves occurring in repeated cycles in RA Discussions. Holmes (1997), based on his research on the discussion sections of social science RAs, found that there is no entirely obligatory Move in this RA section. Similarly, in a study carried out by Peacock (2002), no compulsory Move was found in the corpus of 252 discussion sections from seven disciplines.

Another research carried out on the discussion sections of RAs is that of Ruiying and Allison (2003), who investigated the results, discussion, conclusion, and pedagogical implication sections of 20 applied linguistics English RAs. Their study aimed to explore the relationship between the sections and their communitive purposes. However, they did not examine the potential impact of nativeness and contextual factors, such as the first language writing conventions, on the information organisation, i.e., the realisation of Moves. As in Kanoksilapatham (2003), Ruiying and Allison (2003) used two levels of structure, i.e., Move and Step, in their analysis to "distinguish the communicative purposes from the rhetorical techniques realising the purposes" (p. 379). Rather than a linear scheme, they proposed a hierarchical seven-Move framework for RA Discussions in applied linguistics, further elaborated in the following section. According to the authors, commenting on results and reporting results are respectively obligatory and quasi-obligatory, and the rest are optional Moves in the discussion sections. They further elaborate that the communicative focus in this section is on the obligatory Move, i.e., commenting on results, where the significance of findings is established, and interpretation goes beyond the 'objective' results.

Even though there are some similarities between the rhetorical structure of RAs, previous research has shown that the realisation of Moves may vary in different sections, depending on the genre and field of study. In his analysis of written medical texts, Nwogu (1989) identified two Moves in RA Discussions, comprising several constituent elements. In a later study, Lewin et al. (2001) found five Moves prototypical for RA Discussions in social science, with each Move consisting of one or more component acts. In her rhetorical model for biochemistry RAs, Kanoksilapatham (2003) determined four Moves in Discussions, each realised by various steps. All these Discussion frameworks share particular Moves, such as stating the research outcome, offering interpretations, and indicating the significance of the study. However, while research limitations were indicated in medical (Nwogu, 1989) and biochemistry (Kanoksilapatham, 2003) RA Discussions, they were not outlined by Lewin et al. (2001) in social science texts. Furthermore, the potential counterclaims and implications of results were found to be common Moves in social science and biochemistry RA Discussions, whereas they were not addressed in medical texts.

RA Discussions have been the topic of considerable research, primarily to explore cross-disciplinary variations (e.g., Holmes, 1997; Hopkins & Dudley-Evans, 1988; Peacock, 2002; Ruiying & Allison, 2003; Stoller & Robinson, 2013) and cross-linguistic characteristics of rhetorical structures (Arabi, 2019; Loi et al., 2016; Sadeghi & Alinasab, 2020; Sol-

er-Monreal, 2015). However, few comparative, cross-cultural studies focus on the discourse structures of RAs written by NES and NNES (e.g., Moreno, 2021). Previous research has demonstrated the effect of socio-cultural factors on the written product of non-native speakers (e.g., Ahmadi, 2022; Marefat & Mohammadzadeh, 2013; Moreno, 2021; Tahririan & Jalilifar, 2004). That said, most cross-cultural, genre-based studies have focused on abstract (e.g., Kafes, 2015; Marefat & Mohammadzadeh, 2013; Samar et al., 2014; Tankó, 2017) and introduction sections (e.g., Del Saz Rubio, 2011; Lu et al., 2021; Soler-Monreal, 2015). Little evidence is available about comparative, cross-cultural, genre-based studies on the discussion sections of applied linguistics RAs written by NES and Persian native speakers as NNES, thus calling for further research.

This study fills the gap in previous research by investigating the use of rhetorical Moves in terms of frequency, type, status, structure, and sequence in applied linguistics RA Discussions written by NESs and NNESs. Additionally, attempts were made to analyse significant features from a socio-cultural and contextual perspective. Moreover, the conformance of RA Discussions written by NESs and NNESs to Ruiying and Allison's (2003) Move framework was studied and analysed. In other words, the analysis sought to determine how closely RA Discussions written by NESs and NNESs adhered to the Move framework outlined by Ruiying and Allison (2003) concerning the rhetorical structure and whether nativeness had an impact on the organisation of information. Considering that Ruiying and Allison (2003) developed their framework based on the structural organisation of applied linguistics RAs, regardless of the authors' native language, the present research looked into the consistency of the framework, taking linguistic and contextual factors into account. The results of this study will have pedagogical implications in teaching English for academic (EAP) and specific purposes (ESP), writing practices, and research by assisting instructors in selecting and adapting their teaching materials. As Anthony (2018) points out, the biggest challenge in the classroom is choosing suitable materials. On top of that, the selected teaching materials need to be adapted (Darwis, 2019; Woodrow, 2017), for which the learners' needs should be identified. The findings may also prove helpful to novice NNES academic writers in English. As novice NNESs typically use discursive patterns typical of their native language, publishing their research in English may present additional challenges; hence, the need for them to learn about the rhetorical structure and organisation of information in English RAs.

METHODS

Data Collection

A collection of 150 English RA Discussions in applied linguistics was initially formed, comprising 75 Discussions written by NES and 75 Discussions written by Iranian NNES, published between 2013 and 2020 in international journals. Overall, 80 English RA Discussions, 40 from each corpus, were chosen and compiled. The NNES corpus was selected from the following Scopus-indexed journals: *System, The Journal of AsiaTEFL, Iranian Journal of Language Teaching Research, Journal of Research in Applied Linguistics, Applied Research on English Language.* The NES corpus was selected from the following Scopus-indexed journals: *Applied Linguistics, Journal of English for Academic Purposes, Language Learning & Technology, Language Teaching Research, The Modern Language Journal.*

A simple random sampling method was used to create the corpora to ensure an unbiased sample. The RAs were carefully chosen for this study based on predetermined criteria, and those that did not comply were excluded. First and foremost, the RAs were restricted to empirical studies which conformed with the IMRDC (Introduction-Method-Results-Discussion-Conclusion) standard. In cases where functional headings such as "Results and Discussions" and "Discussion and Conclusion" were employed, the RAs with a distinct subsection and subheading for Discussion were selected; otherwise, they were ignored. As for the authors' nationality, judgements were primarily made based on their names; when in doubt, an online search for their background, affiliation, and native language was conducted. The RAs were excluded when the authors' nationality could not be verified, a paper by the same corresponding author had already been selected, and a non-native-speaking scholar was involved in a multi-authored RA.

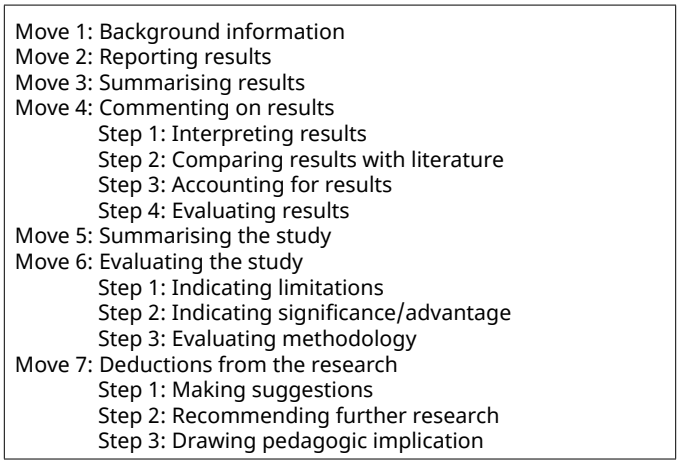
Data Analysis

Both qualitative and quantitative methods were applied to conduct the rhetorical Move analysis of RA Discussions. For the identification of Moves and Steps, the hierarchical seven-Move framework of Ruiying and Allison (2003) was employed since it was developed based on research on applied linguistics which is the focus of the present study. Ruiying and Allison's (2003) model is a two-layer analysis comprising Moves and Steps, where a Move is realised by one or more Steps (Figure 1). A Move "captures the function and purpose of a segment of text at a more general level" (Ruiying & Allison, 2003, p. 370) and can involve one or more Steps, defined as rhetorical realisations of the function of Move.

To identify Moves and Steps in texts, they were segmented into sentences as units of analysis. Sentences are considered a complete unit of meaning and have been commonly used in the genre analysis literature (e.g., Basturkmen, 2012; Holmes, 1997; Zhang & Wannaruk, 2016). Besides, since there were very few instances of writers embedding two or more moves within the same sentence, adopting a unit of analysis below that level was deemed unnecessary. Having divided the texts into the unit of analysis, each sentence was carefully examined and annotated by Move and Step based on its communicative intent. In cases where a unit of text served multiple communicative functions, it was assigned

Figure 1

Ruiying and Allison's (2003) seven-Move model for the research article Discussions



to the most salient Move and Step (e.g., Del Saz Rubio, 2011; Holmes, 1997). The results of the initial qualitative analysis of Moves and Steps were subsequently subjected to the quantitative analysis. More specifically, the frequency and percentage of Moves and Steps in each RA Discussion and across all RAs in each corpus, along with the percentage of RAs per corpus featuring each Move and Step, were calculated and recorded. The obtained results were tabulated and analysed per corpus to determine the structural pattern and complexity of Moves. Additionally, statistical tests were conducted to determine whether the rhetorical structure of RA Discussions differed significantly between the two corpora.

The quantitative results were compared using the Mann-Whitney *U* test for statistical variations in Move occurrence between the two datasets. The test was conducted on the relative frequency, i.e., the percentage of each Move against the total frequency of Moves per corpus, to eliminate the impact of text length variation. The status of Moves and Steps in each corpus was subsequently identified based on their relative frequency. The cut-off point for Move and Step classification was set based on Kanoksilapatham's (2005) criterion, classifying each Move as obligatory, conventional, or optional in terms of their occurrence in 100%, 60-99%, and below 60% of RAs, respectively. Chunks of text were selected as the unit of analysis to evaluate the cyclicity of Moves in Discussions according to their communicative intent. To elaborate, a group of adjacent sentences featuring the same Move constituted one chunk of text. Text chunks were annotated, tabulated, and analysed for their structure and cyclicity. The present study conducted a literature search to identify the contextual factors that may have influenced the use of Moves by NNES writers.

Reliability Measures

The two corpora were analysed and coded on two occasions within a 2-month interval to establish the reliability of Move identification. The intra-rater reliability was measured using Cohen's kappa test for each corpus, both of which were calculated to be above 0.95.

RESULTS AND DISCUSSION

Rhetorical Move Analysis

To investigate the employment of Moves and Steps in terms of their frequency, status, structure, and sequence between RA Discussions in NES and NNES corpora, the relative frequency of Moves and Steps was juxtaposed to disregard the effect of text length variation. The Mann-Whitney *U* tests showed no statistically significant difference in the occurrence of Move 1, i.e., background information ($U = 700.5$, $p = 0.336$), as well as Move 2, i.e., reporting results ($U = 609$, $p = 0.065$), between the NES corpus ($f_{DM1} = 16.39\%$, $f_{DM2} = 20.51\%$) and the NNES corpus ($f_{DM1} = 19.40\%$, $f_{DM2} = 21.40\%$). However, a significant difference was observed in the occurrence of Move 3, i.e., summarising results ($U = 468$, $p = 0.001$) between the two corpora – the NES corpus ($f_{DM3} = 4.76\%$) and the NNES corpus ($f_{DM3} = 2.31\%$). As for Move 4, even though there was no significant difference in the overall occurrence of the Move ($U = 779.5$, $p = 0.843$) between the NES corpus ($f_{DM4} = 38.19\%$) and the NNES corpus ($f_{DM4} = 51.5\%$), significant differences were observed in the second Step, i.e., comparing results with literature ($U = 587$, $p = 0.04$), and the fourth Step, i.e., evaluating results ($U = 456.5$, $p = 0.01$). The tests showed no statistically significant difference in the occurrence of Move 5, i.e., summarising the study ($U = 781$, $p = 0.663$), which was barely used in the discussion sections of both the NES corpus ($f_{DM5} = 0.24\%$) and the NNES corpus ($f_{DM5} = 0.31\%$). The most significant difference was observed in Move 6, i.e., evaluating the study ($U = 427$, $p = 0$) between the NES corpus ($f_{DM6} = 11.5\%$) and the NNES corpus ($f_{DM6} = 3.23\%$). All the three Steps of Move 6 had significant differences – the largest of which was observed in the first Step, i.e., indicating

limitations (U = 492.5, p = 0.001), accounting for 5.64% and 0.92% of the NES corpus and the NNES corpus, respectively. Similarly, the test showed a statistically significant difference in Move 7, i.e., deductions from the research, and all the three Steps of the Move (U = 447, p = 0) between the NES corpus (f_{DM7} = 8.87%) and the NNES corpus (f_{DM7} = 2%). The largest difference in Move 7 was observed in the first Step, i.e., making suggestions (U = 487, p = 0), comprising 3.88% and 0.87% of the NES corpus and the NNES corpus, respectively. To further elaborate, Table 1 summarises the frequency of Moves and Steps across all RAs.

As shown in Table 1, Move 4 was identified as the most frequent Move in both corpora, which was an expected outcome since it is the communicative purpose of Discussions to elaborate, establish meaning, and indicate the contribution of findings in the field of study. This result is consistent with that of Ruiying and Allison (2003). However, Move 4 was employed significantly more by NNESs, who dedicated more than half of the Discussions to this Move. With regard to the Steps, there was a significantly higher tendency in NNES RAs to compare research results with previous studies. In contrast, NES RAs approached this Move primarily by justifying the research results. A closer examination revealed that NNESs also relied on earlier studies for reasoning purposes. This finding may have cultural underpinnings. It is possible that in the Iranian context, authors are more concerned about the validity of their judgment and, thus, are more inclined to seek approval for their opinions and interpretations in the light of other studies. Correspondingly, in his research on Iranian master's theses, Nodoushan (2011) refers to Iranian writers' concern with the validity of their research. Nevertheless, this hypothesis will need further qualitative investigation.

Regarding the Steps of Move 4, as the most common Move in both corpora, some discrepancies were observed between the findings of this study and those of Ruiying and Allison (2003). In their research, the most prominent Steps across the applied linguistics RA Discussions were "interpreting results" and "accounting for results". However, although the latter was the most frequent Step employed by NESs and the second most frequent Step used by NNESs in this study, "interpreting results" is the least frequent Step in both corpora. The results of the present study indicated "comparing results with literature" to be the most prominent Step in the realisation of "commenting on results" employed by NNESs and the second most frequent Step used by NESs, which is in contrast with Ruiying and Allison's (2003) findings, as well. Furthermore, contrary to their finding that Move 1, i.e., background information, was rarely used in Discussions, the current study found that both NESs and NNESs employed this Move extensively, making it the third most common Move in Discussions. In addition to Move 1, Ruiying and Allison (2003) found Move 5, i.e., summarising the study, equally infrequent in RA Discussion. In the same vein, the results of the current study showed this Move to be the least frequent Move in both corpora, indicating that it may not be expected in Discussions.

To determine the status of each Move and Step in RA Discussions, in terms of obligatory, conventional, or optional, the frequency and percentage of RA Discussions featuring each Move and Step were calculated. Table 2 presents the status of Moves and Steps in both NES and NNES corpora. The results showed similarities and differences in Move occurrence in the two data sets. Move 1 (background information), Move 2 (reporting results), and Move 4 (commenting on results) were found to be the obligatory Moves across all RA Discussions in both NES and NNES corpora, indicating their inclusion by all the authors. This finding is, to some extent, inconsistent with that of Ruiying and Allison (2003), showing "commenting on results" as the only obligatory Move in RA Discussions, with "reporting on results" being quasi-obligatory, i.e., employed in all but one RA Discussions. Similar findings were reported by Basturkmen (2012) and Le and Harrington (2015), indicating "commenting on results" to be the most crucial Move in the discussion sections in dentistry and applied linguistics RAs, respectively. Moreover, reporting research results, i.e., Move 2, has been reported as an obligatory Move in several previous studies, as well (e.g., Alamri, 2020; Amirian et al., 2008; Atai & Falah, 2005; Dudley-Evans, 1994; Hopkins & Dudley-Evans, 1988; Nodoushan, 2011; Peng, 1987). However, little research has reported Move 1 as an obligatory Move (e.g., Rasmeenin, 2006, as cited in Nodoushan, 2011).

Move 6, i.e., evaluating the study, and Move 7, i.e., deductions from the research, were identified as conventional Moves in the NES corpus, having occurred in most RA Discussions. In contrast, the results indicated that these two Moves were optional in the NNES corpus, with less than one-third of the RAs employing them. This finding, which is consistent with the results of other studies (e.g., Arabi, 2019; Atai & Falah, 2005; Ruiying & Allison, 2003), might be due to the reason that the appearance of these Moves in Discussions is influenced by the subsequent sections, such as Conclusion (Ruiying & Allison, 2003). Accordingly, in RAs that include both sections, there seems to be a tendency to comment on results in Discussions and evaluate the study in Conclusions. The lack of tendency of NNESs to assess their findings may also be attributed to a cultural assumption in the Iranian context that evaluation is a task for others, not one's own. Undoubtedly, this hypothesis will require further investigation from cultural and contextual perspectives.

Move Sequence and Structure

Regarding the sequence, pattern, and cyclicity of Moves, it was observed that the NESs and NNESs did not necessarily progress sequentially through Moves and Steps; that is, RA Discussions in both corpora had a highly cyclical structure. This finding supports previous research highlighting cyclic patterns, rather than linear or compositional, i.e.,

Table 1

Frequency and Percentage of Occurrence of Moves and Steps

Move/Step		NES (N = 40)			NNES (N = 40)		
		f	*%f*	$\bar{x}$	*f*	*%f*	$\bar{x}$
Move 1	Background information	279	16.39	6.98	252	19.40	6.30
Move 2	Reporting results	349	20.51	8.73	278	21.40	6.95
Move 3	Summarising results	81	4.76	2.03	30	2.31	0.75
Move 4	Commenting on results	650	38.19	16.25	667	51.35	16.68
Step 1	Interpreting results	62	3.64	1.55	76	5.85	1.90
Step 2	Comparing results with literature	204	11.99	5.10	297	22.86	7.43
Step 3	Accounting for results	217	12.75	5.43	201	15.47	5.03
Step 4	Evaluating results	167	9.81	4.18	93	7.16	2.33
Move 5	Summarising the study	4	0.24	0.10	4	0.31	0.10
Move 6	Evaluating the study	188	11.05	4.70	42	3.23	1.05
Step 1	Indicating limitations	96	5.64	2.40	12	0.92	0.30
Step 2	Indicating significance/advantage	50	2.94	1.25	14	1.08	0.35
Step 3	Evaluating methodology	42	2.47	1.05	16	1.23	0.40
Move 7	Deductions from the research	151	8.87	3.78	26	2.00	0.65
Step 1	Making suggestions	66	3.88	1.65	7	0.54	0.18
Step 2	Recommending further research	51	3.00	1.28	8	0.62	0.20
Step 3	Drawing pedagogic implications	34	2.00	0.85	11	0.85	0.28

Note. f = frequency of Move/Step across all RAs. %f = the percentage of Move/Step across all RAs.
$\bar{x}$ = average occurrence of Move/Step per section.

the compositions of Moves without adhering to the linear or cyclic structure, in RA Discussions (e.g., Amnuai & Wannaruk, 2013; Atai & Falah, 2005; Holmes, 1997; Hopkins & Dudley-Evans, 1988; Loi et al., 2016; Peacock, 2002; Peng, 1987; Swales, 1990). Peacock (2002), for example, highlights the presence of Move cycles in RA Discussions, particularly those written by NNESs specialising in language and linguistics. In the present study, the analysis of the sequence of Moves demonstrated the highly cyclical structure of Discussions in the majority (96%) of RAs in the NNES corpus (e.g., M1-M2-M4-M1-M2-M4-M2-M4-M2-M4-M2-M4-M6-M4-M7), and all the RAs in the NES corpus (e.g., M4-M6-M2-M4-M6-M3-M4-M2-M4-M6-M1-M6-M4-M6). This finding supports Peacock's (2002) notion regarding the highly cyclical structure of Moves in the NNES corpus.

The most frequent two-Move cycle in both corpora was Move 2, i.e., reporting results, followed by Move 4, i.e., commenting on results, an example of which is presented below. Among all RA Discussions featuring two-Move cycles, 38% in the NNES corpus and 23% in the NES corpus were realised through the repetition of Move 2 followed by Move 4. In other words, NNESs showed a higher tendency to comment on the results immediately after their report, which, as mentioned before, was mainly done by comparing their findings with previous studies. In contrast, NESs approached this Move primarily by accounting for the results. The prevalence of Move 2 instead of Move 3, i.e., summarising the results, indicates that both groups of NES and NNES writers tend to establish their evaluations based on individual, main results rather than on a projection of findings.

Furthermore, the complexity of the cyclical structure of RA Discussions was analysed according to Move repetitions. The results showed that 100% of NES RAs and 90% of NNES RAs employed one-Move and two-Move cycles. At a more complex level, 67.5% of NES RAs and 45% of NNES RAs featured three-Move cycles in their Discussions (e.g., **M3-M4-M2**-M4-M6-M2-**M3-M4-M2-M3-M4-M2-**...). Only one instance of the four-Move cycle was seen in NES RAs (**M1-M3-M2-M4**-M1-M2-M4-M2-M4-M6-**M1-M3-M2-M4**-...), with no occurrences in the NNES corpus. These findings indicated a more complex cyclical pattern in NES RA Discussions, which may have been due to the employment of more Move categories in the RAs, as shown in Table 1 and Table 2.

An example of the cyclicity of Moves:

> *[Move 2:] The results of the study [...] revealed that the participants' written accuracy in revised compositions was significantly affected*

by OIF. [Move 4/Step 3:] This finding is in line with the findings of the study conducted by Nassaji (2011), who explored the role of oral negotiation in response to written errors in L2 classrooms. […] [Move 2:] Given the findings of the second research question, the results of the study indicated that OIF had a significant effect on the written accuracy of the participants on the post-test in the OIF group in comparison with the EF group. [Move 4/Step 3:] With respect to OIF and retention, this finding supports the results of the study by Lyster and Saito (2010b), who investigated the impact of different kinds of oral CF on learners' oral errors and found that CF plays a facilitative role for L2 development and that its impact is sustained at least until delayed posttests. (NNES, D6)

Apart from the structural pattern, the opening and closing Moves were also studied. The Moves featured in NES and NNES RAs to open and conclude Discussions are presented in Table 3. As illustrated, there is a similarity between the two corpora regarding the main opening Move. Most RAs in both NES and NNES corpora tended to start Discussions with Move 1 (background information). However, Move 3 (summarising the results) was more frequently used as the opening Move in NES RAs compared to their NNES counterparts, where it was used only once for this purpose. Furthermore, the two corpora differed significantly in their featured Moves to conclude Discussions. NESs showed an almost equal tendency to close the section with Move 7 (deductions from the research) or Move 4 (commenting on results), mainly to introduce further research and evaluate the results, respectively. However, a significant proportion of NNES RAs concluded Discussions using Move 4, most often by comparing the results with previous studies. Contrary to NES RAs, Move 7 (deductions from the research) and Move 6 (evaluating the study) were notably less used in NNES RAs to conclude Discussions. A closer surface-level examination of NNES RAs showed that non-native writers tended to employ these Moves in RA Conclusions rather than Discussions.

Examples of opening Moves:

[Move 1:] The purpose of the present study was to examine the significance of teachers' self-efficacy and collective teacher efficacy in predicting teachers' psychological well-being among Iranian English teachers. (NNES, D2)

[Move 3:] This study has demonstrated that analytic judgments of collocation accuracy, lexical diversity, and word frequency are highly predictive of holistic judgments of lexical proficiency for both written and spoken samples. (NES, D16)

Table 2

Distribution and Status of Moves and Steps

Move/Step		NES (N = 40)			NNES (N = 40)		
		n	%	Status	n	%	Status
Move 1	Background information	40	100	***	40	100	***
Move 2	Reporting results	40	100	***	40	100	***
Move 3	Summarising results	29	72.5	**	16	40	*
Move 4	Commenting on results	40	100	***	40	100	***
Step 1	Interpreting results	26	65	**	30	75	**
Step 2	Comparing results with literature	37	92.5	**	36	90	**
Step 3	Accounting for results	38	95	**	33	82.5	**
Step 4	Evaluating results	39	97.5	**	32	80	**
Move 5	Summarising the study	3	7.5	*	2	5	*
Move 6	Evaluating the study	30	75	**	19	47.5	*
Step 1	Indicating limitations	21	52.5	*	8	20	*
Step 2	Indicating significance/advantage	20	50	*	9	22.5	*
Step 3	Evaluating methodology	17	42.5	*	5	12.5	*
Move 7	Deductions from the research	26	65	**	12	30	*
Step 1	Making suggestions	19	47.5	*	4	10	*
Step 2	Recommending further research	15	37.5	*	6	15	*
Step 3	Drawing pedagogic implications	16	40	*	7	17.5	*

Note. n = the number of RAs featuring Move/Step; % = the percentage of RAs featuring Move/Step.
**** = obligatory.*
*** = conventional.*
** = optional.*

Figure 2

Typical cyclic patterns of the core Moves

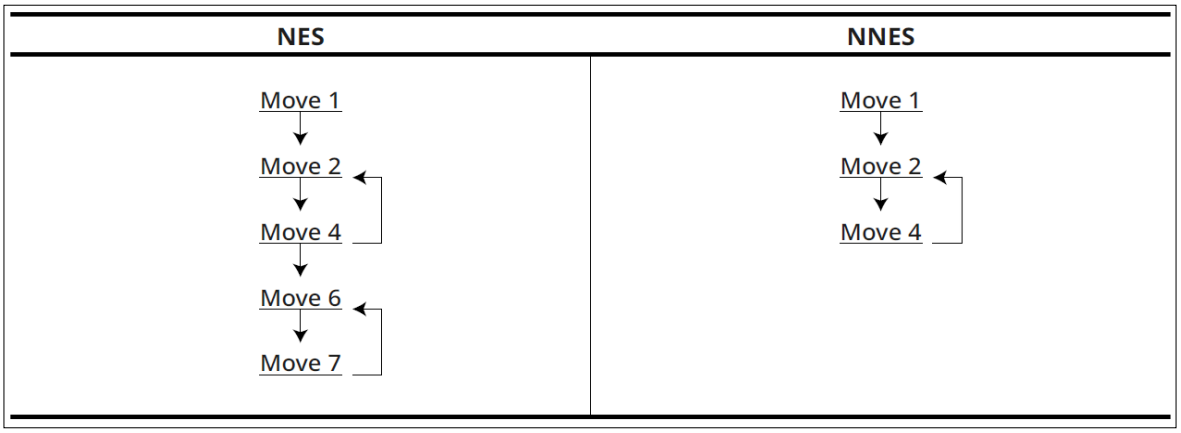


Table 3

Distribution of Opening and Closing Moves

	Opening Moves			Closing Moves		
	Move	**n**	**%**	**Move**	**n**	**%**
NES	Move 1	24	60	Move 7	16	40
(N = 40)	Move 3	8	20	Move 4	14	35
	Move 2	5	12.5	Move 6	6	15
	Move 4	3	7.5	Move 2	4	10
NNES	Move 1	30	75	Move 4	23	57.5%
(N = 40)	Move 2	6	15	Move 7	7	17.5%
	Move 4	2	5	Move 2	5	12.5%
	Move 3	1	2.5	Move 1	3	7.5%
	Move 6	1	2.5	Move 6	2	5.0%

Examples of closing Moves:

> *[Move 4/Step 2:] As mentioned earlier, the literature has documented a positive relationship between teachers' individual efficacy and collective efficacy, which hand in hand influence student achievement (Goddard & Goddard, 2001; Goddard et al., 2000, Goddard, Hoy, & Woolfolk Hoy, 2004; Kurz & Knight, 2004; Tschannen-Moran & Barr, 2004). (NNES, D33)*

> *[Move 7/Step 1:] Though neither the survey findings nor the interview findings are generalisable, they nevertheless suggest a need for greater attention to diversity in sampling for studies of proofreading in higher education. (NES, D34)*

Overall, the results indicated conformance with Ruiying and Allison's (2003) Move framework concerning the types of Moves. However, no linear pattern of Moves and Steps was found in any Discussions. Moreover, the results did not support the chosen framework regarding the status of Moves. This disagreement in results may be due to the distinction between the datasets used for the studies. In Ruiying and Allison's (2003) research, only 8 RAs with a distinct discussion section were included in the corpus of study. However, the present research meticulously selected 80 RAs that followed the IMRDC structure. The reliability of findings may be affected by the sample size.

The results of this study showed parallels and variances in the organisation of information, i.e., realisation of Moves, between NES and NNES corpora. Considering the relative frequency of Moves, the most prominent, i.e., obligatory, Moves across each corpus were similar, i.e., Move 1, Move 2, and Move 4, with some differences in their level of employment (see Tables 1 and 2). On the other hand, there were significant disagreements between the two sets of data regarding the Steps of obligatory Moves and the employment of conventional and optional Moves. A cyclic structure of Moves was found in both NES and NNES corpora, featuring the same two-Move cycle, i.e., Move 2-Move4. Nevertheless, the differences in the occurrence of Moves across the two corpora were also evident in their cyclic patterns. As illustrated, Moves 6 and 7 were used cyclically by NESs but not by

their NNES counterparts (see Figure 2). Furthermore, NESs and NNESs mainly used the same Move to open the Discussion, i.e., Move 1; however, their strategy differed for closing the section (see Table 3). Researchers have previously attributed writers' rhetorical choices to cultural and social contexts (e.g., Kafes, 2015; Marefat & Mohammadzadeh, 2013). Previous research has also indicated the effect of cultural and contextual conventions on the written product of Iranian researchers (e.g., Marefat & Mohammadzadeh, 2013; Tahririan & Jalilifar, 2004). Past writing experiences, writing strategies in first and second language (Zhang, 2018), and fear of negative evaluation (Schmidt, 2004; Shang, 2013), have also been identified as influential factors in second language writing. As discussed earlier, the differences observed between the two sets of RA Discussions in this study might be due to any of these issues.

It is essential to acknowledge some limitations in the current study to avoid over-generalising the findings. The present research focused on English RAs in one academic discipline, i.e., applied linguistics, and analysed the rhetorical Move structures in one section of published RAs, i.e., Discussions. Future research may expand upon the inclusion of different academic disciplines and other sections of RAs. Moreover, the study's data was limited by size; thus, the generalizability of the current findings can be increased by expanding the size of the datasets. It is also recommended that future studies include the investigation of lexico-grammatical features, i.e., linguistic realisations, of the rhetorical Moves in NES and NNES RAs. Furthermore, interviews can be conducted with RA authors to analyse the impact of their educational background, whether they have been taught academic writing, cultural background, and other influential factors on their written products, as explained above.

CONCLUSION

The present comparative study provides an insight into the distinct communicative intents and functions of RA Discussions written by NESs and NNESs, i.e., Iranian scholars. For non-native English-speaking scholars to maintain their professional standing nationally and internationally, and in response to the so-called academic doctrine of 'publish or perish', it is imperative to write and publish RAs in English, which is the academic lingua franca in the global context. The findings of this study provide further research-based evidence that the rhetorical organisation and structure of RAs are culturally and contextually dependent.

The present research offers several practical and pedagogical implications in the context of English for academic or specific purposes, particularly for research publication purposes. Researchers can benefit from such analytical comparisons by gaining a clearer understanding of the differences in norms and practices of scholarly communication in local and international discourse communities. Novice researchers, in particular, can gain valuable insight into the common rhetorical practices in Anglophone academic writing and the effect of first-language contextual factors. Such knowledge and awareness will enable non-native English-speaking scholars to make more informed rhetorical choices and to present their arguments more effectively. Furthermore, education specialists could find the results helpful in preparing appropriate materials and designing curriculums tailored to the learners' needs and aligned with the rhetorical conventions adopted by the academic community. Move analysis can also help teachers become more familiar with the distribution of information across RA Discussions. Teachers can use this research-based knowledge to set more practical objectives, provide better feedback, adopt appropriate instructional strategies, and design tasks and assignments that facilitate novice writers' understanding of rhetorical patterns.

CONFLICT OF INTERESTS

The author declares that she has no conflict of interest.

REFERENCES

Ahmadi, L. (2022). Interactional Metadiscourse Markers in Scientific Texts (Based on Research Articles Written by Native and Non-Native Speakers). *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2. Yazykoznanie* [Science Journal of Volgograd State University. Linguistics], *21*(4), 99-110. https://dx.doi.org/10.15688/jvolsu2.2022.4.7

Alamri, B. (2020). A comparative study of Saudi and international journals of Applied Linguistics: The move–bundle connection approach. *Journal of Language and Education*, *6*(2), 9-30. https://dx.doi.org/10.17323/jle.2020.10531

Alinasab, M., Gholami, J., & Mohammadnia, Z. (2021). Genre-based revising strategies of graduate students in applied linguistics: Insights from term papers. *Journal of English for Academic Purposes*, *49*, 100941. https://dx.doi.org/10.1016/j.jeap.2020.100941

Amirian, Z., Kassaian, Z., & Tavakoli, M. (2008). Genre analysis: An investigation of the discussion sections of applied linguistics research articles. *The Asian ESP Journal*, *4*(1), 39-63.

Amnuai, W. (2019). The textual organization of the discussion sections of accounting research articles. *Kasetsart Journal of Social Sciences*, *40*(2), 389-394.

Amnuai, W., & Wannaruk, A. (2013). Investigating Move Structure of English Applied Linguistics Research Article Discussions Published in International and Thai Journals. *English Language Teaching*, *6*(2), 1-13. https://dx.doi.org/10.5539/elt.v6n2p1

Anthony, L. (2018). *Introducing English for specific purposes*. Routledge.

Arabi, H. A. (2019). Move analysis of the discussion section in Arabic and English research articles. *ESP Across Cultures*, *16*, 9-25.

Atai, M. R., & Falah, S. (2005). A contrastive genre analysis of result and discussion sections of applied linguistic research articles written by native and non-native English speakers with respect to evaluated entities and ascribed values. 10th Conference of Pan-Pacific Association of Applied Linguistics, Edinburgh, United Kingdom.

Basturkmen, H. (2012). A genre-based investigation of discussion sections of research articles in Dentistry and disciplinary variation. *Journal of English for Academic Purposes*, *11*(2), 134-144. https://dx.doi.org/10.1016/j.jeap.2011.10.004

Brett, P. (1994). A genre analysis of the results section of sociology articles. *English for Specific Purposes*, *13*(1), 47-59. https://dx.doi.org/10.1016/0889-4906(94)90024-8

Burrough-Boenisch, J. (2003). Shapers of published NNS research articles. *Journal of second language writing*, *12*(3), 223-243. https://dx.doi.org/10.1016/S1060-3743(03)00037-7

Chen, T.-Y., & Kuo, C.-H. (2012). A Genre-Based Analysis of the Information Structure of Master's Theses in Applied Linguistics. *The Asian ESP Journal*, *8*(1), 24-52.

Cho, S. (2004). Challenges of Entering Discourse Communities Through Publishing in English: Perspectives of Nonnative-Speaking Doctoral Students in the United States of America. *Journal of Language, Identity & Education*, *3*(1), 47-72. https://dx.doi.org/10.1207/s15327701jlie0301_3

Cross, C., & Oppenheim, C. (2006). A genre analysis of scientific abstracts. *Journal of Documentation*, *62*(4), 428-446. https://dx.doi.org/10.1108/00220410610700953

Darwis, N. (2019). Adapting In-Service Teacher Training Materials and Activities into Teaching English as a Foreign Language Materials as a Studi Program Lecture (A Classroom Action Research at the Group 3, 6th Semester English Education Study Program of STAIN Watampone). *Ekspose: Jurnal Penelitian Hukum dan Pendidikan*, *17*(2), 694-708. https://dx.doi.org/10.30863/ekspose.v17i2.125

Del Saz Rubio, M. M. (2011). A pragmatic approach to the macro-structure and metadiscoursal features of research article introductions in the field of Agricultural Sciences. *English for Specific Purposes*, *30*(4), 258-271. https://dx.doi.org/10.1016/j.esp.2011.03.002

Dudley-Evans, T. (1994). Genre analysis: An approach to text analysis for ESP. In M. Coulthard (Ed.), *Advances in Written Text Analysis* (pp. 219-228). Routledge.

Glasman-deal, H. (2020). *Science Research Writing: For Native And Non-native Speakers Of English (Second Edition)*. World Scientific Publishing Company.

Hatzitheodorou, A.-M. (2014). A Genre-oriented Analysis of Research Article Abstracts in Law and Business Journals: Anna-Maria Hatzitheodorou. In M. Bondi & R. L. Sanz (Eds.), *Abstracts in Academic Discourse* (pp. 175-198). Peter Lang. https://dx.doi.org/10.3726/978-3-0351-0701-2/19

Holmes, R. (1997). Genre analysis, and the social sciences: An investigation of the structure of research article discussion sections in three disciplines. *English for Specific Purposes*, *16*(4), 321-337.

Hopkins, A., & Dudley-Evans, T. (1988). A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes*, *7*(2), 113-121. https://dx.doi.org/10.1016/0889-4906(88)90029-4

Hu, G., & Liu, Y. (2018). Three minute thesis presentations as an academic genre: A cross-disciplinary study of genre moves. *Journal of English for Academic Purposes*, *35*, 16-30. https://dx.doi.org/10.1016/j.jeap.2018.06.004

Kafes, H. (2015). Cultural or Discoursal Proclivity: Rhetorical Structure of English and Turkish Research Article Abstracts. *The Anthropologist*, *21*(1-2), 240-254. https://dx.doi.org/10.1080/09720073.2015.11891813

Kanoksilapatham, B. (2003). *A corpus-based investigation of biochemistry research articles: Linking move analysis with multidimensional analysis* [Doctoral Thesis, Georgetown University].

Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, *24*(3), 269-292. https://dx.doi.org/10.1016/j.esp.2004.08.003

Le, T. N. P., & Harrington, M. (2015). Phraseology used to comment on results in the Discussion section of applied linguistics quantitative research articles. *English for Specific Purposes*, *39*, 45-61. https://dx.doi.org/b10.1016/j.esp.2015.03.003

Lei, J., & Hu, G. (2019). Doctoral candidates' dual role as student and expert scholarly writer: An activity theory perspective. *English for Specific Purposes*, *54*, 62-74. https://dx.doi.org/10.1016/j.esp.2018.12.003

Lewin, B. A., Fine, J., & Young, L. (2001). *Expository Discourse: A Genre-based Approach to Social Science Research Texts*. Continuum.

Lim, J. M. H. (2006). Method sections of management research articles: A pedagogically motivated qualitative study. *English for Specific Purposes*, *25*(3), 282-309. https://dx.doi.org/10.1016/j.esp.2005.07.001

Loi, C. K., Evans, M. S., Lim, J. M.-H., & Akkakoson, S. (2016). A Comparison Between Malay and English Research Article Discussions:A Move Analysis. *SAGE Open*, *6*(2), 1-11. https://dx.doi.org/10.1177/2158244016652925

Lu, X., Yoon, J., & Kisselev, O. (2021). Matching phrase-frames to rhetorical moves in social science research article introductions. *English for Specific Purposes*, *61*, 63-83. https://dx.doi.org/10.1016/j.esp.2020.10.001

Marefat, H., & Mohammadzadeh, S. (2013). Genre analysis of literature research article abstracts: A cross-linguistic, cross-cultural study. *Applied Research on English Language*, *2*(2), 37-50. https://dx.doi.org/10.22108/are.2013.15469

Moreno, A. I. (2021). Selling research in RA discussion sections through English and Spanish: An intercultural rhetoric approach. *English for Specific Purposes*, *63*, 1-17. https://dx.doi.org/10.1016/j.esp.2021.02.002

Nodoushan, M. (2011). A structural move analysis of discussion sub-genre in applied linguistics. 2012: 6th International Conference on Languages, E-Learning and Romanian Studies, Isle of Marstrand, Sweden.

Nwogu, K. N. (1989). *Discourse variation in medical texts : Schema, theme and cohesion in professional and journalistic accounts* [Doctoral Thesis, Aston University].

Peacock, M. (2002). Communicative moves in the discussion section of research articles. *System*, *30*(4), 479-497. https://dx.doi.org/10.1016/S0346-251X(02)00050-7

Peng, J. (1987). Organizational features in chemical engineering research articles. *English Language Research Journal*, *1*(4), 79-116.

Ruiying, Y., & Allison, D. (2003). Research articles in applied linguistics: Moving from results to conclusions. *English for Specific Purposes*, *22*(4), 365-385. https://dx.doi.org/10.1016/S0889-4906(02)00026-1

Sadeghi, K., & Alinasab, M. (2020). Academic conflict in Applied Linguistics research article discussions: The case of native and non-native writers. *English for Specific Purposes*, *59*, 17-28. https://dx.doi.org/10.1016/j.esp.2020.03.001

Sadeghi, K., & Shirzad Khajepasha, A. (2015). Thesis writing challenges for non-native MA students. *Research in Post-Compulsory Education*, *20*(3), 357-373. https://dx.doi.org/10.1080/13596748.2015.1063808

Samar, R. G., Talebzadeh, H., Kiany, G., & Akbari, R. (2014). Moves and steps to sell a paper: a cross-cultural genre analysis of applied linguistics conference abstracts. *Text & Talk*, *34*, 759-785.

Samraj, B. (2002). Introductions in research articles: variations across disciplines. *English for Specific Purposes*, *21*(1), 1-17. https://dx.doi.org/10.1016/S0889-4906(00)00023-5

Santos, M. B. D. (1996). The textual organization of research paper abstracts in applied linguistics. *Text - Interdisciplinary Journal for the Study of Discourse*, *16*(4), 481-500. https://dx.doi.org/10.1515/text.1.1996.16.4.481

Schmidt, L. A. (2004). Psychometric evaluation of the writing-to-learn attitude survey. *Journal of Nursing Education*, *43*(10), 458-465. https://dx.doi.org/10.3928/01484834-20041001-04

Shang, H.-F. (2013). Factors associated with English as a foreign language university students writing anxiety. *International Journal of English Language Teaching*, *1*(1), 1-12.

Soler-Monreal, C. (2015). Announcing one's work in PhD theses in computer science: A comparison of Move 3 in literature reviews written in English L1, English L2 and Spanish L1. *English for Specific Purposes*, *40*, 27-41. https://dx.doi.org/10.1016/j.esp.2015.07.004

Stoller, F. L., & Robinson, M. S. (2013). Chemistry journal articles: An interdisciplinary approach to move analysis with pedagogical aims. *English for Specific Purposes*, *32*(1), 45-57. https://dx.doi.org/10.1016/j.esp.2012.09.001

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Swales, J. (2004). *Research Genres: Explorations and Applications*. Cambridge University Press. https://dx.doi.org/10.1017/CBO9781139524827

Tahririan, M. H., & Jalilifar, A. R. (2004). Generic analysis of thesis and dissertation abstracts: variation across cultures. *Iranian Journal of Applied Linguistics*, *7*(2), 121-143.

Tankó, G. (2017). Literary research article abstracts: An analysis of rhetorical moves and their linguistic realizations. *Journal of English for Academic Purposes*, *27*, 42-55. https://dx.doi.org/10.1016/j.jeap.2017.04.003

Woodrow, L. (2017). *Introducing course design in English for specific purposes*. Routledge.

Yang, W. (2016). Evaluative language and interactive discourse in journal article highlights. *English for Specific Purposes*, *42*, 89-103. https://dx.doi.org/10.1016/j.esp.2016.01.001

Zhang, B., & Wannaruk, A. (2016). Rhetorical Structure of Education Research Article Methods Sections. *PASAA: Journal of Language Teaching and Learning in Thailand*, *51*, 155-184. https://dx.doi.org/10.14456/pasaa.2016.6

Zhang, Y. (2018). Exploring EFL Learners' Self-Efficacy in Academic Writing Based on Process-Genre Approach. *English Language Teaching*, *11*(6), 115-124. https://dx.doi.org/10.5539/elt.v11n6p115

https://doi.org/10.17323/jle.2022.11917

Teachers' Perceptions of Promoting Student-Centred Learning Environment: An Exploratory Study of Teachers' Behaviours in the Saudi EFL Context

Mohammad Al-khresheh

Northern Border University, Arar, Saudi Arabia

ABSTRACT

Background. Although the constructivist and humanistic theories advocate for a more student-centred learning approach, the contemporary prac**tice** of English teachers is more oriented towards the behavioural approach in the Saudi EFL context.

Purpose. This study aimed to explore teachers' perceptions of promoting student-centred learning in the Saudi EFL context. It initially amplifies elements of effective student-centred EFL teaching and merges the elements with teachers' behaviours, resulting in four measurable categories: assessment strategy, communicative approach to learning, teachers' qualifications, and group activities in teaching EFL.

Methods. Following the quantitative approach, a questionnaire on teachers' observable behaviours, comprising 47 items and categorised under four constructs, with a three-point Likert scale, was carried out on a group of 302 English teachers. The instrument was devised through an online survey. To analyse data, descriptive and inferential statistical analysis was used. Exploratory factor analysis was performed to validate the study's instrument.

Results. The interpretation of data primarily showed teachers' self-reported practices were more oriented toward a teacher-centred approach, specifically in teaching grammar. Other findings showed teachers' positive attitude towards student-centred learning in lesson preparation, using ICT to prepare interactive teaching materials and giving constructive feedback. Teachers had a mediocre attitude towards their professional qualifications, indicating that the teachers might need more effective English language training workshops based on their needs. Additional data also proved that there was no significant difference in teachers' perception in terms of their gender, location and status of the school, and classroom size.

Implications. This study provides an overview of the teachers' position on integrating student-centred learning in their English classroom and thus accentuates the need for potential opportunities for the teachers' professional development and demands less content overload so that the teachers can ample their teaching process appropriately.

KEYWORDS

Constructivism, English, foreign language teaching, humanistic approach, perceptions, teaching behaviours, Saudi EFL Context, student centred-learning

Citation: Al-khresheh, M. (2022). Teachers' Perceptions of Promoting Student-Centred Learning Environment: An Exploratory Study of Teachers' Behaviours in the Saudi EFL Context. *Journal of Language and Education, 8*(3), 23-39. https://jle.hse.ru/article/view/11917

Correspondence:
Mohammad Al-khresheh
mohd_khresheh@yahoo.com

Received: January 06, 2021
Accepted: September 22, 2022
Published: September 30, 2022

INTRODUCTION

English is a foreign language (EFL) that has a tremendous role in the Kingdom of Saudi Arabia (Al-khresheh, 2020a). Aside from its importance for educational purposes, it also has a crucial role in accomplishing the objectives of Saudi Vision 2030 (Al-Zahrani & Rajab, 2017). Additionally, it is widely acknowledged that to develop a new breed of global citizens (United Nations, 2015), the targeted development goals must also include learning an FL; therefore, EFL usage assumes even more significance (Almazova et al., 2016).

Despite this, many studies often report that teaching EFL effectively tends to be difficult as far as Saudi Arabia is concerned (Al-khresheh, 2020b; Al-Seghayer, 2014; Al-Zahrani & Rajab, 2017). No

doubt, this is a highly concerning issue that can only be resolved with a dedicated policy. Additional research must also be carried out to determine the reasons for this problem, along with possible solutions. This study is focused on tackling these issues that come with effective EFL teaching.

Education experts have called for student-centred language learning environments to be established to resolve multiple difficulties and promote the effective teaching of the English language in teaching-learning activities. Several countries, including Saudi Arabia, currently use a teacher-centred learning style, which involves instructing or lecturing students repeatedly to teach English. Using this style indicates that the Grammar-Translation method (GMT) is still alive, especially in teaching grammar. The teacher-centred forms-focused approach can result from the domination of the GTM where the main objective is to understate discrete rules and produce only correct forms (Alfares, 2017; Assalahi, 2013; Kassem, 2018; Mohammad & Hazarika, 2016). However, this approach is in disagreement with the constructivist point of view and appears to have been inspired by the behaviourist point of view. This perspective has been criticised repeatedly in several studies. Additionally, when this approach is used to teach EFL, students feel pressured into memorising their course syllabus, which cripples their autonomy, practice, and motivation to use the knowledge learned in the classroom practice in their daily lives (Song & Kim, 2017).

For the promotion of effective student-centred learning environments to teach EFL, the most widely accepted and cited studies today recommend the following: (a) integration of practical large or small group activities, discussions, or works to improve interest and motivation levels among students learning a FL; (b) constant communication between teachers and their students, between the students themselves, and between students and society at large, to boost student participation levels; (c) offering immediate feedback on a real-time basis as part of assessments and to do away with pressure-inducing external examinations, which are infamous for triggering anxiety issues and diminishing the learning abilities of students; (d) the establishment of a special teachers' qualification that prioritises the autonomy of the learners and defines the role of the teachers as facilitators, not instructors (Al-Zahrani & Rajab, 2017; Alfares, 2017; Almazova et al., 2016; Rind et al., 2019; Sidhu et al., 2020).

Four of the attributes mentioned above in this study regarding student-centred learning and teaching activities have been prioritised. These attributes are used along with results from a previous study to measure FL teachers' behaviour (Bell, 2005)- assessment strategy, communicative approach to learning, teachers' qualifications, and group activities in teaching EFL. These four attributes also feed the ten elements of efficient student-centred teaching-learning of English as a foreign language (Jacobs & Renandya, 2016). This study further provides a conceptualisation of combining efficient student-centred learning of foreign language and measuring English teachers' observable behaviours as depicted in Info-graph 1.

The existing literature also shed light on several harsh realities that were preventing teachers from teaching EFL effectively, such as a high-class size (Yi, 2008), content overload (Roussel et al., 2017), and external examinations (Baksh et al., 2016). Simultaneously, this literature also emphasised that teachers' perceptions could overcome these challenges and lower anxiety levels among students learning a second language (Dewaele et al., 2018).

Teacher perception is generally dependent on the following factors: their primary level of perceived competence, self-efficacy (Al-khresheh, 2020c; You et al., 2019), the level of influence their beliefs have over their teaching strategies (Debreli, 2016), their teaching experience, their learning process (Bonner et al., 2019), continuous development programs (You et al., 2019), and their expectations regarding the outcome and efficiency of their teaching strategies (Wozney et al., 2006). This study defines EFL teachers' perspective as a dynamic combination of several attitudes and beliefs they display, which can be observed through their classroom behaviours (Bell, 2005).

Given that teacher perception has a huge role in the constructive learning of EFL among students, this study intends to research the perception of language teachers about various salient features inherent to a student-centred learning and teaching environment and how an efficient language teacher might affect them. These salient features have been determined from existing literature by evaluating various effective student-centred learning practices in an English language teaching environment (Jacobs & Renandya, 2016) and the observable behaviours that reveal teacher perception regarding effective strategies used to teach EFL (Bell, 2005). The existing literature has yielded four major attributes - continuous assessment, communicative approaches, teacher qualities, and group activities in teaching English.

Furthermore, this study also seeks to determine the influence wielded by factors like class size, content overload and external examination on teachers' perceptions of effective language teaching, which can be exposed through teachers' behaviours.

To accomplish the research goals laid out in this study, the following questions are posed: (1) How do teachers perceive their role as effective teachers in addressing student-centred learning in the Saudi EFL context? In light of this main research question, this study aims to find out more about the EFL teachers' behaviour in the following aspects: assessment strategy, communicative approach to learning, teachers' qualifications, and group activities in teaching EFL. (2) How do different variables (i.e., gender, class size, school status, and location) influence teachers' behaviours towards

promoting a student-centred learning environment in their class in the Saudi EFL context?

LITERATURE REVIEW

Student-Centred Learning: A Revolutionary Paradigm Shift from Behaviourism to Constructivism - Humanism

Behaviourist education theory tended to be the primary approach used to analyse the students' acquisition of knowledge beyond the mental process, rather a psychological and educational process and dominated the educational field during the 1920-1950s (Watson, 2017). This theory considers knowledge to be something prevailing outside a person, acquired through behaviour modification. Learning is a behavioural modification that may be conditioned appropriately using positive or negative reinforcement measures such as rewards or punishments (Phillips, 2012). Accordingly, students are rendered as responders more willingly than actors. Their responses to the stimuli are subject to evaluation. If the response happens to be undesirable, it is the teacher's responsibility to offer appropriate feedback that reprograms students to have a more appropriate response in the future. The teacher shall also ensure that undesirable responses from the students are discouraged in future evaluations (Devaki, 2021; Orak & Al-khresheh, 2021; Wubante, 2020). However, advancing educational theories tended to evolve from a behavioural perspective considering this approach solely focuses on acquiring new conceptual knowledge. The behaviourist approach to education was criticised for being too rigid and instruction-based (Schunk, 2012) and teacher-centred in several studies (Wilson & Myers, 2000). Educators reported that the behaviourist approach had tremendous negative repercussions and lowered interest levels (Pritchard, 2017), resulting in the memorisation of the content knowledge without real-life implications.

With a particular focus on student-centred learning, the cognitivist theory approach recognised learning as the assimilation of prior knowledge while simultaneously accommodating new information, resulting in the transfer of knowledge (Pritchard, 2017). This theory addressed the students as something beyond blank slates and acknowledged the importance of their prior knowledge as a foundation for constructing new behaviours or imparting new knowledge. Nevertheless, the cognitivist theory was criticised by contemporary educators for overlooking the importance of reinforcements that offered motivation to the learners (Overskeid, 2008).

Based on the cognitivist approach, constructivism theory expanded the scope of learning. It contextualised the learners' personal experiences and views in the actual learning process to provide students with a sense of relevance and thus provided a satisfactory resolution to the issue of sustaining their motivation. The constructivist theory of learning recognised learning as a social, situated and metacognitive process in which interaction occurs between prior knowledge and learning outcomes (Aljohani, 2017). Social constructivism theory, in particular, aligns with Vygotsky's sociocultural theory, which emphasises the role of sociocultural context as being a major influence in the learning process (Taylor & MacKenney, 2008). Social constructivism theory further advocates the need for promoting students' active interaction through problem-solving, group decision making or peer discourse measures and programs (Svinicki, 2004). This was a major shift from the instruction-based rigid teacher-centred approach towards a student-centred learning approach by encouraging students' participation in the classroom, thus providing them with the autonomy to take charge of their learning (Deci & Ryan, 2010).

To address the need to focus more on various non-cognitive variables of learning, i.e. students' need, emotions, values and self-perceptions (Snowman et al., 2012), the humanistic education theory emphasises concepts like self-efficacy and self-actualisation among students (Schunk, 2012). It also seeks to improve learners' affective and social skills through cooperative and collaborative learning programs (Huitt, 2009). It also sought to promote critical thinking, bring about positive changes in attitude, encourage logical and persuasive action, and boost social engagement among learners (Tobolowsky, 2014). Constructivists advocate for a student-centred learning program along the lines of Deci and Ryan's self-determination theory, which enables learners to possess an intrinsic motivation to learn and study conceptual knowledge on their own (Ryan & Deci, 2016).

From the information presented above, a conclusion can be drawn that there was a gradual shift in the dominant and widely-recognised learning theories of education towards a student-centred learning approach, which promoted active student participation, communication, and collaboration. This paradigm shift also heralded an evolution of the teachers' role from just issuing instructions or providing/assessing knowledge to being a facilitator of learning and education for their students through continuous assessment, thus encouraging students to have more autonomy and say in their learning process.

Maintaining a Student-Centred Approach in the Effective EFL Teaching

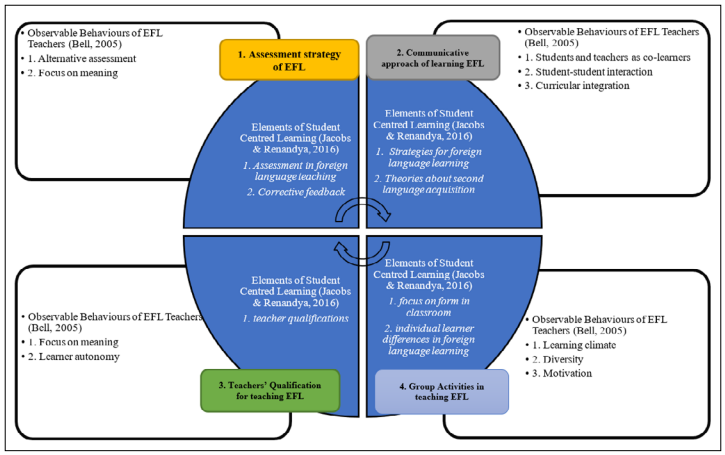
When it comes to teaching EFL, not only are teachers required to act as active facilitators for learning but also to create a student-centred learning environment in which students are allowed to take charge of their learning process. In language teaching, Jacobs and Renandya (2016) presented ten elements of effective student-centred learning, aside from motivation, which included interaction, autonomy, focus on meaning, curricular integration, diversity, learning

climate, alternative assessment, and thinking skills. On the other hand, teacher perception was considered a combination of the teachers' beliefs and attitudes (Alsalem, 2020), which had a major influence on the effective teaching of EFL. This perception is measured in a previous study by researching the teachers' observable behaviours and teaching theories (Bell, 2005). Teachers' observable behaviour refers to the teachers' effectiveness and capability of teaching EFL.

This study conceptualises four major attributes from the abovementioned aspects: sets of observable behaviour and student-centred learning elements. The conceptualisation is visually represented in the following info-graph:

Info-Graph 1

Conceptualising Four Major Attributes for Measuring EFL Teachers' Observable Behaviour Concerning Student-centred Learning



These four major measurable attributes have been explained comprehensively below, along with a justification for the need to have additional research studies on this topic.

EFL Assessment Strategy

All over the world, the evaluation of the students' grasp of EFL is done with external examinations. The same is true in the Saudi Arabian context as well. As a matter of fact, research literature points out that the traditional external examination system used to evaluate one's English language abilities has a negative impact on students since students tend to memorise the course content to pass their examinations (Mohammad & Hazarika, 2016). This results from a behaviourist learning approach, which has been criticised widely for being too teacher-centred, rigid, and material-based (Watson, 2017). To promote the creation of student-centred learning environments in English language teaching in Saudi Arabia, educators point to a gap in existing literature regarding the assessment of EFL (Picard, 2018). As a result, the present study explores teachers' perceptions of various assessment techniques for teaching EFL.

Considering Vygotsky's sociocultural theory of education (Pathan et al., 2018), continuous assessment, including students' self-evaluation, peer evaluation, immediate feedback from teachers, help the learners to form their knowledge constructively while simultaneously helping them take charge of their learning which is the key to student-centred learning. To promote such a student-centred assessment in Saudi Arabia, educators have called for a shift in the teachers' perception of assessment techniques, encouraging them to move from external examination-oriented assessment toward a more accurate classroom assessment strategy (Aldawood, 2016; Alsalem, 2020).

The Communicative Language Teaching Approach

When it comes to learning an FL, students and teachers' interaction assumes critical significance as this encourages students to practice their newly-developed language (Sert, 2019). In addition to student interaction with other students and teachers, students' autonomy also happens to be one of the main factors influencing the learning of an FL. It enables the students to take charge of their learning process (Godwin-Jones, 2019; Murray & Lamb, 2017). Vygotsky's theory of the zone of proximal development advocated for student autonomy, considering it helps students reach the zone of their proximal development while learning a new language. Curricular integration offers learners the opportunity to apply their newly-developed language skills within their language classroom and across and beyond the context of their classrooms (Shabani et al., 2010).

The Saudi Ministry of Education has recognised the need to reform and modernise the strategies used for teaching EFL. To promote communicative competence among students, it has started modifying the curriculum and enabling teachers to embrace various communicative strategies (Abahussain, 2016). However, multiple studies reveal that the teacher-centred classroom learning process and the absence of adequate communicative practice while learning EFL continue to be a reality (Abahussain, 2016; Picard, 2018). This is

why one of the goals of the current study is to explore teachers' perceptions in promoting the communicative approach to teaching EFL.

Teachers' Qualification for Teaching EFL

Teacher qualification is considered to be a broader terminology. It is more like an umbrella that incorporates various attributes of an effective teacher. However, current literature believes teacher quality is the product of teacher training, involvement in teacher development programmes, and in-service or pre-service courses (Köksal & Ulum, 2019; Wilden et al., 2020). In the Saudi EFL context, the qualifications possessed by these teachers to teach EFL are often questionable (Al-Hazmi, 2003; Picard, 2018). Al-Zahrani and Rajab's study (2017) revealed that teachers' perception regarding the qualification they possessed for teaching English was not very convincing. As a result, one of the objectives of this current study is to explore the teachers' perceptions regarding their qualifications to teach EFL.

Group Activities in Teaching EFL

From a constructivist perspective, in line with Vygotsky's sociocultural theory, group activities enable students to conceptualise and internalise the learning content with self-actualisation and self-determination (Deci & Ryan, 2010; Huitt, 2009). In order to learn English effectively, worldwide educators advocate for group activities to achieve the learning objectives of the four crucial skills of language teaching - reading, writing, speaking and listening (A-Garni & Almuhammadi, 2019; Alrayah, 2018; Ochoa et al., 2016).

However, in the Saudi EFL context, the current literature reveals that though most students prefer group activities to help them learn English, the teachers prefer to fall in line with the traditional teacher-oriented learning approach (Alfares, 2017; Alnujaidi, 2019). This is why one of the primary objectives of this study is to explore the teachers' perceptions regarding group activities' effectiveness for teaching EFL.

Current Scenario of EFL in the Saudi Context

Alrabai (2016), in his study, drew particular emphasis on the abysmally low success rates recorded among students who were learning EFL in the context of Saudi Arabia. This research sheds light on a few possible underlying factors, categorised as internal and external factors. External factors or individual factors, which impact the students' ability to learn EFL, are identified as – using Arabic as the first language for education, sociocultural issues, the role of religion, influence of a teacher, teaching method, curriculum, and education system as a whole. The internal factors are gender, demography, attitude, aptitude, motivation, language learning strategies, anxiety, and learners' autonomy in learning EFL.

Al-Awaid (2018) identified similar factors behind the low success rates of EFL. However, they emphasised and diagnosed problems cropping up at the classroom level instead of blaming on policy failures. The insights revealed by this study showed that students had a dismal lack of interest in using English, and their teachers did not possess adequate training to address the situation. However, this study provides numerous recommendations to improve the current state of things by encouraging communicative practices in the classroom and real-life application of EFL, establishing strategic planning measures to overcome challenges like content overload, large classrooms, and so on.

Alharbi's review of the current state of teaching EFL (2019) gives an overview of the history of EFL in Saudi Arabia and how the teaching training programs and curricula have evolved. The study also emphasised other literature findings, which revealed current EFL training programs to be 'teacher-centred classrooms by the GTM' and advocated for a more student-centred approach. The study's findings align with the conclusions laid out by previous studies while investigating the mitigating factors of learning EFL. Besides, this study also suggests that teachers favour increased engagement. Yet, they find English language difficult, which has become a substantial barrier to meaningful progress.

The current study assumes more significance as it combines two pieces of literature to explore the teachers' perceptions regarding student-centred learning through four major attributes and by measuring the behaviours a group of EFL teachers exhibits in the Saudi EFL context.

METHODS

The study's primary objective was to explore the teachers' perceptions of four different attributes of effective student-centred learning through sets of observable EFL teachers' behaviour. The study follows a quantitative research approach to achieve its main objectives because the variables, which the study aims to explore, are measurable numerically (Bell, 2005; Creswell, 2012; Jacobs & Renandya, 2016).

Participants

In most quantitative survey research methods, researchers select a sample size that accurately represents the population after considering their inability to use the whole population as the sample (Tolmie et al., 2011). For this study, a total of 302 Saudi EFL teachers were selected to participate in the survey. The sampling technique was followed by purposeful sampling since the sampling was dependent on the participants' willingness to contribute to the study (Palinkas et al., 2015). The distribution of the sample was approximately equal. The variations of the participant teachers' profiles have been explained in Table 1 below:

Table 1
Distribution of the Sample

Gender		School location		School status		Class size		
Male	Female	Urban	Rural	Public	Private	≤ 30	30-50	≥50
47%	53%	47%	53%	51%	49%	13%	57%	30%

Instrument

As explained in the literature review section, the instrument's theoretical construct covers two crucial areas: effective student-centred learning and observable behaviours exhibited by EFL teachers that encourage effective language teaching. Four major attributes were identified from the two areas. Table 2 offers a brief overview of the constructs of the instrument.

Table 2
Conceptualisation of the Instrument

EFL teachers' observable behaviours (Bell, 2005)	Four attributes derived from literature	Elements of student-centred learning (Jacobs & Renandya, 2016)
Assessment in FL teaching (8 items)	assessment strategy	Alternative assessment
Corrective feedback (7 items)		Focus on meaning
FL learning strategies (4 items)	Communicative approach	Students and teachers as co-learners
Theories about second language acquisition (5 items)		Student-student interaction
		Curricular integration
Teacher qualifications (7 items)	Teachers' qualifications	Focus on meaning
		Learner autonomy
Focus on form in classroom (10 items)	Group activities	Learning climate
Individual learner differences in FL learning (6 items)		Diversity
		Motivation

The instrument was chosen from the pre-existing literature (Bell, 2005). The instrument consists of 47 items. These 47 items appeared on the questionnaire randomly, not consistent with category. Participants were assured that their responses would be kept confidential. The original instrument used a 5-point Likert scale. However, in this study, after due consideration of the fact that the participant number was much smaller than in the original study, a 3-point Likert scale was adopted (He et al., 2020). The Likert scale was interpreted as - Disagree (value 1), Neutral (value 2), Agree (value 3).

Data Collection and Analysis

A targeted email list of EFL teachers was formulated through extensive research. An email campaign was set up using MailChimp and Getresponse to deploy an email campaign/bulk messages to the generated contact list with a pre-designed email template, which contained a link to a Google Forms survey embedded in it as a call-to-action button. Willing participants responded to the email, and the survey was controlled when the target participant number was reached. To analyse data, Exploratory Factor Analysis was carried out to validate the constructs of the instrument. The results revealed relevant constructs, as mentioned in the previous instrument. Correlation was measured to determine if there was any repetition of items, but the results showed no repetition. All items' loading was higher than 0.40, which is an acceptable load. JASP was used to carry out a descriptive and split case analysis. Besides, the final outputs were consolidated in Microsoft Excel to build figures and charts.

Validity and Reliability of the Instrument

The validity was established already in the published instrument. Additionally, Exploratory Factor Analysis showed the same constructs. The reliability was measured by Cronbach Alpha (Peters, 2018), the overall score of which was 0.906, indicating the instrument was reliable.

RESULTS

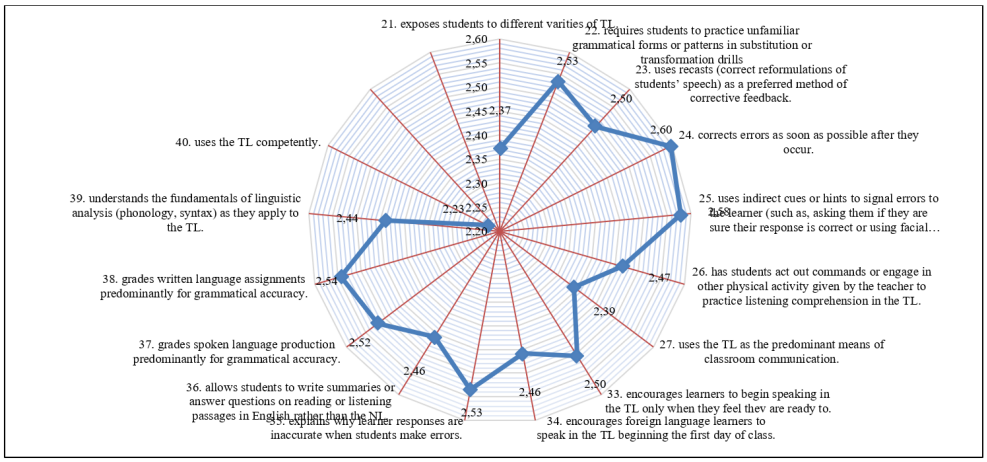
In this section, the data analysis of this study is presented based on the research questions. The study's preliminary results (table of item statistics and figure of overall responses and overall mean score) have been given in Appendix 1.

Teachers' Behaviours Regarding the Need for Continuous EFL Assessments in the Saudi EFL Context

Figure 1 below represents EFL teachers' behaviours towards addressing continuous assessment in the Saudi EFL context (EFL Assessment, corrective feedback).

Figure 1

Teachers' Behaviours towards Addressing Continuous EFL Assessment



The abovementioned figure shows the participants' high priority on providing immediate feedback to students as soon as they make a mistake in their language usage. This immediate feedback may be through verbal or nonverbal means of communication (items 24, 25). The participant EFL teachers also favour linguistic analysis to be applied to the target language (TL) (item 39). However, as the figure reveals, the participants are often unable to use the TL effectively (item 40).

In general, from the mean score analysis, it is pretty evident that the teachers had a positive attitude in general towards the usage of continuous EFL assessments (least value more than 2.2 out of 3 maximum). Nevertheless, the item analysis also reveals that the teachers were slightly less positive towards TL usage as the predominant method to teach EFL or even the introduction of different dialects of TL to their students.

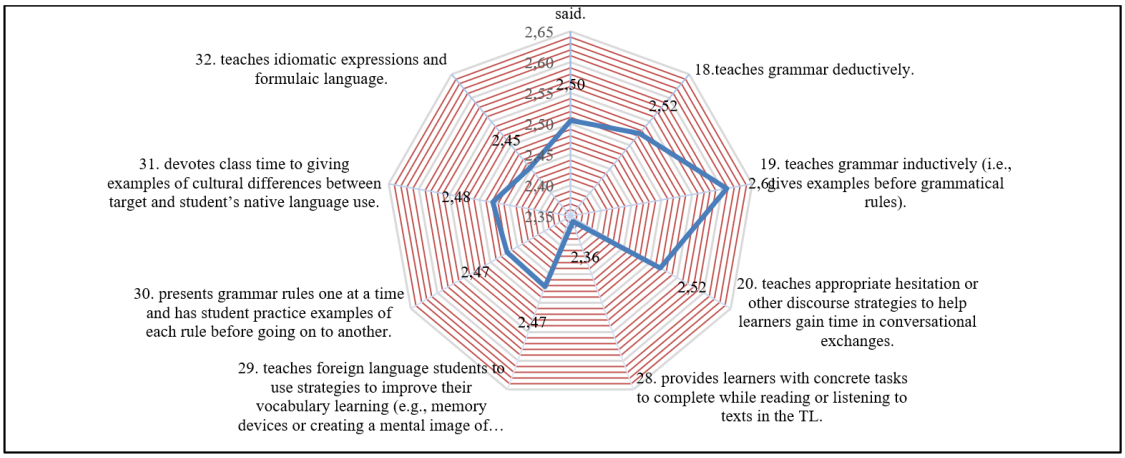
When it came to determining the most effective student-centred approach to teaching EFL, a continuous assessment was used as an alternative assessment tool. The response of teachers towards alternative assessment through corrective feedback was rather positive. However, the impact on the student-centred approach was relatively low.

Teachers' Behaviours towards the Communicative Approach for Learning EFL

Figure 2 represents the behaviours of EFL teachers towards using a communicative approach for learning EFL in the Saudi context (Strategies for FL learning and theories about second language acquisition).

Figure 2

EFL Teachers' Behaviours towards Using the Communicative Approach



The figure above reveals that this study's participants heavily emphasised curriculum integration. Additionally, the participant teachers prioritised the inductive method for teaching English grammar (item 19), followed by thoroughly explaining the grammar rules before asking students to practice the relevant structure. This attribute reflects the communication between teachers and students, primarily in the participant teachers' classrooms. However, the participant teachers were also compassionate towards their students' understanding of the TL (item 17).

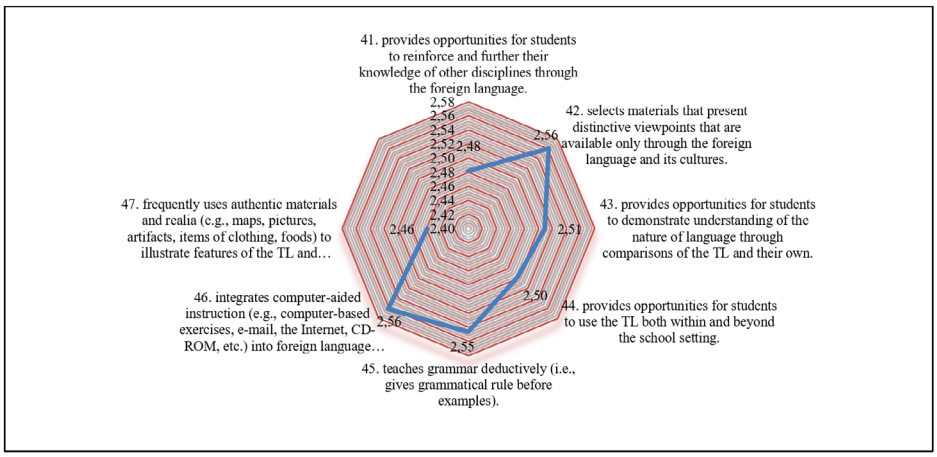
The concept of an effective student-centred learning environment requires teachers to appreciate and treat students like co-learners. The lowest teachers' perception was towards allowing students to complete a concrete task while listening and reading the TL (item 28); indicating that they did not value student autonomy in the classroom. Nevertheless, teachers had a relatively higher perception of promoting the communicative approach for effective student-centred teaching (the least mean value was 2.35 with a maximum of 2.65).

EFL Teachers' Behaviours regarding Qualifications for the Accomplishment of Learning Objectives

Figure 3 represents the behaviours of EFL teachers regarding qualifications for the accomplishment of learning objectives (teachers' qualifications).

Figure 3

EFL teachers' Behaviours towards their Qualifications to Achieve Learning Objectives



The above figure paints a good picture regarding teachers' behaviour in student-centred teaching of EFL. The mean score was highest in this attribute. Participant teachers displayed a higher perception regarding technology-based exercises and even selected materials from the English language culture (item 42, 46). However, the participant teachers were found to have a less favourable opinion regarding everyday items as teaching aids (item 47) even though there was ample evidence to suggest that this technique could be advantageous in schools that lacked adequate resources.

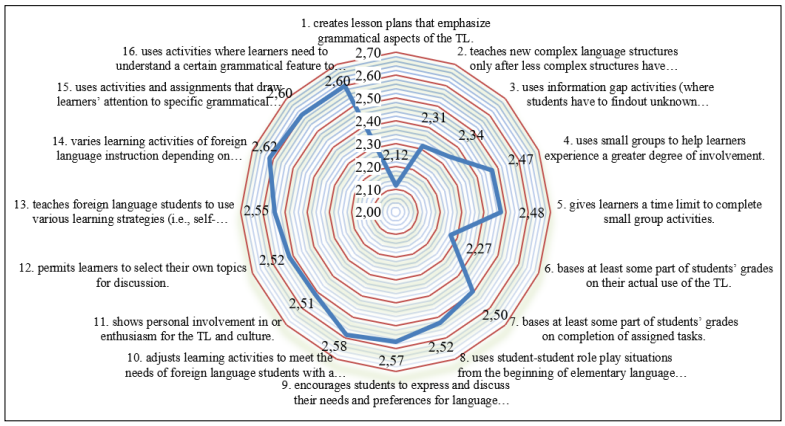
Teacher quality is important in establishing a student-centred learning environment for teaching EFL. In this case, the findings reveal that teachers' ability to handle student-centred learning required appropriate training.

Teachers' Behaviours towards Using Group Activities for Effective EFL Learning

The following figure represents teachers' behaviours towards using group activities for effective learning of EFL (focus on form in classroom, individual learner differences in EFL learning).

Figure 4

Teachers' Behaviours towards Using Group Activities for Effective EFL Learning



This figure reflects that teachers acknowledge the diversity of their student population and consider it while teaching EFL (item 14). The participant teachers also actively promoted group activities, small group tasks, and assignments for properly disseminating the English language, especially regarding grammar. The teachers were less favourable when drafting strategic lesson plans for their students (item 1).

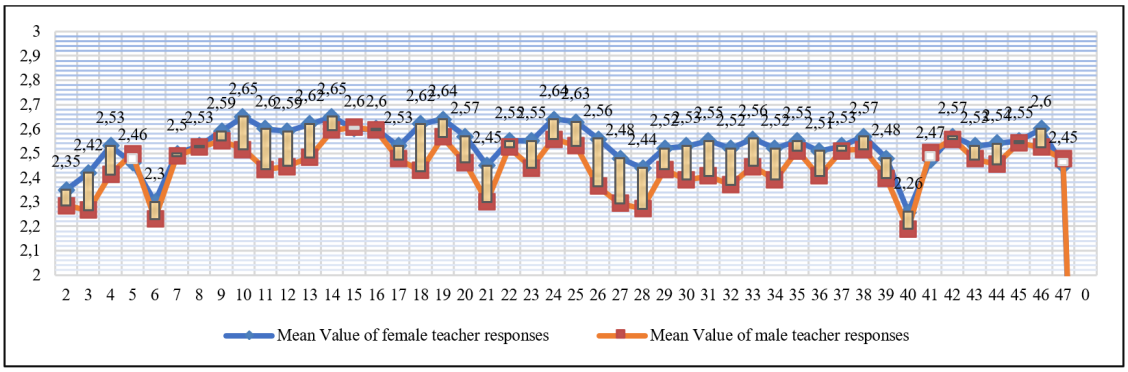
Given that the student-centred learning approach places a heavy emphasis on the students' motivation, diversity, and self-determination, the insights revealed by this figure show that the participant English teachers were quite open to the idea of engaging students in student-centred learning. Although the teachers did not welcome the integration of group activities into their language assessment tests (item 6), the data showed that participant teachers were open to the idea of group activities to teach EFL.

The influence of Gender, School Status, School Location, and Class Size on EFL Teacher's Behaviours towards Using Student-centred Approach

An independent-samples t-test statistical method was sued to figure out the difference. It was found that there are no statistically significant differences in teachers' perceptions in terms of their gender, school location (urban or rural), and school status (public or private). A one-way ANOVA test was used to determine the difference according to the class size variable (less than 30, having 30-50, or higher than 50). No significant difference was also found. Figure 5 represents EFL teachers' perspectives on using a student-centred approach based on their gender.

Figure 5

The Influence of EFL Teachers' Gender on their Behaviours for the Promotion of Student-centred Learning

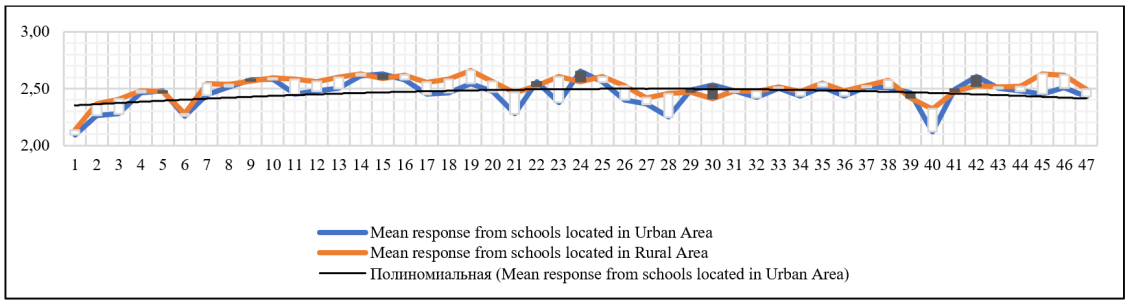


This figure reveals that there is not much variation in male and female teachers' behaviours which are encouraged to promote effective student-centred learning. However, in general, the female teachers positively responded towards promoting the four major attributes of effective student-centred learning. On the other hand, the male teachers failed to prioritise the promotion of corrective behavioural aspects and consider student diversity aspects.

Figure 6 below represents the influence of school location on teachers' Behaviours towards using a student-centred approach.

Figure 6

The Influence of School Location on EFL Teachers' Behaviours towards Promoting Student-centred Approach

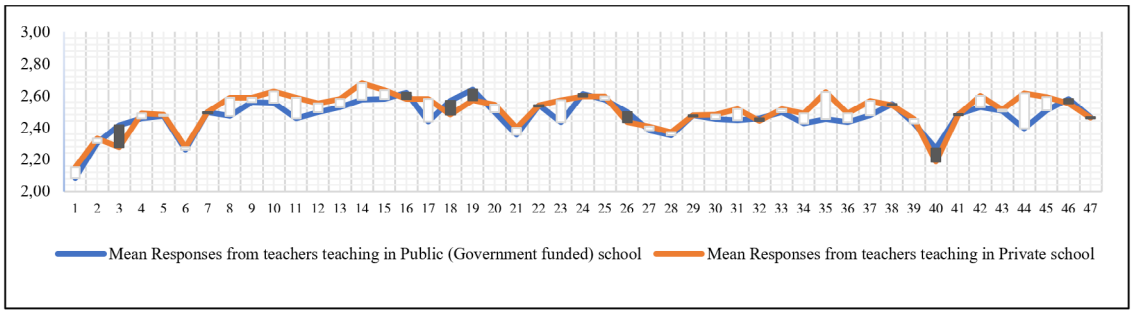


The above figure shows that urban schools tilt towards a student-centred environment more than rural schools. The participant English language teachers who hailed from urban areas responded positively towards integrating the four crucial attributes of effective student-centred learning compared to their counterparts from rural areas.

Figure 7 below represents school status's influence on teachers' behaviours towards using a student-centred approach.

Figure 7

The Influence of School Status on EFL Teachers' Behaviours towards Promoting Student-centred Approach

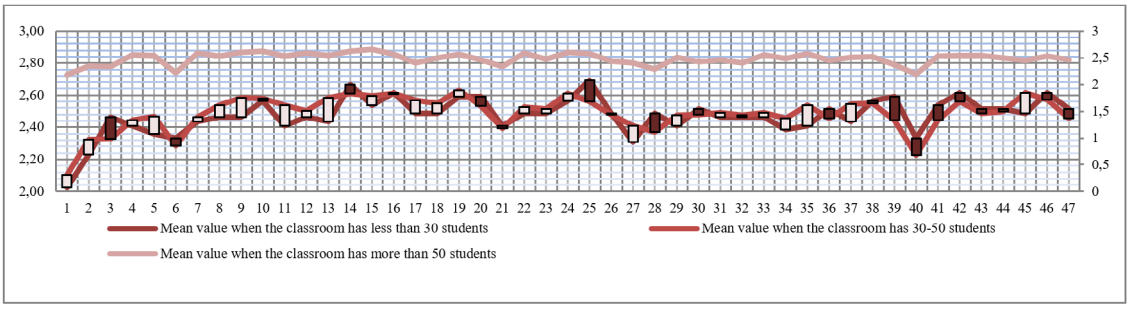


In general, the abovementioned figure reveals that the school's status has a minimal impact on teachers' behaviour to promote student-centred learning in teaching EFL. In most cases, the participant teachers from public schools had a statistically insignificant higher positive response. However, statistically speaking, the difference is negligible, as mentioned earlier.

Like the school's status, the class size also had minimal influence on teachers' behaviours towards promoting student-centred learning in EFL, as depicted in figure 8 below. However, in some cases, classes with a student count of more than 50 were found to have participant teachers who responded positively towards the idea of promoting group activities, tasks, and assignments compared to classes with 30-50 students or less than 30 students. Nevertheless, the statistical value also happens to be negligible in this case.

Figure 8

The influence of Class-size on EFL Teachers' Behaviours towards Promoting Student-centred Approach



DISCUSSION

The study seeks to understand the teachers' perceptions regarding the concept of effective student-centred learning by measuring their behaviours while teaching EFL with an emphasis on four major attributes in the Saudi EFL context, namely, EFL assessment strategy, the communicative approach to learning, teachers' qualifications, and group activities in teaching EFL (see Table 2). Moreover, the influence of variables such as gender, class size, school status and location on teachers' behaviours towards promoting student-centred approach in the Saudi EFL context have also been put under the lens in this study.

Although the thorough analysis of data shows that the existing teaching-learning method is a bit more teacher-centred practice than a student-centred one while teaching grammar as an example, the overall data gathered from the participants generally paints a positive picture of teachers' behaviours towards incorporating a student-centred approach in their classrooms.

On further analysis of the four major attributes, some concerning aspects have emerged out of this study. The first attribute is the EFL Assessment strategy. Continuous assessment practices are a major concern for student-centred learning environments (Aldawood, 2016). A promising result from this study revealed that the participant teachers responded positively to the idea of correcting errors made by students immediately when they occur, either through verbal or non-verbal communication methods (Figure 1, items 24-25). This reflects their position on immediate performance feedback for students actively engaged in-class activities, a traditional classroom practice. However, this fails to bolster the case for continuous EFL assessment since student participation may vary significantly depending on the circumstances. Students have higher anxiety levels while learning a foreign language, which also depresses their classroom participation levels (Alrabai, 2016). While academically-proficient students have higher participation rates in classroom activities, which give teachers more opportunities to correct their errors, low-performing students have lower participation rates. They do not have as many opportunities to be corrected by their teachers if they make errors. Another noteworthy aspect was the discouraging response the study's participants gave regarding their ability to use the TL competently (Figure 1, item 40). The possible reason behind teachers' failure to complete their TL may lay in content overload, which refers to the completion of a curriculum part within a time frame and their lack of strategic planning of time management. This finding was not particularly surprising as previous studies have talked extensively about the issue of content overload in the EFL curriculum, acting as a barrier to the proper dissemination of EFL teaching (Alrabai, 2016). However, strategic planning and proper time management practices are highly recommended by existing literature for EFL teachers looking to complete their TL effectively in the Saudi EFL context (Al-Awaid, 2018).

Reflecting on the usage of a communicative approach in English classrooms for encouraging student-centred learning, teachers' responses to items (1, 6, 21, 28, and 40) show that the existing teaching-learning method is a more teacher-centred practice than a student-centred one while teaching grammar. Their responses to these items confirm that they lack priority in teaching English grammar with a constructive and interactive approach. However, the participant teachers were also open to the idea of explaining grammar using an inductive method that involved the usage of examples in explaining grammar rules (Figure 2, item 19). The participants also provided a contradictory responses when issuing concrete tasks and goals to students learning to read and write (Figure 2, item 28). This response indicates that the participant teachers did not hold a favourable view regarding the importance of learners' autonomy in language learning. Learners' autonomy is considered critical in encouraging a student-centred EFL teaching approach (Godwin-Jones, 2019; Murray & Lamb, 2017). The traits that are responsible for contradictory positions assumed by these teachers can be traced back to the traditional behaviourist approach, where the teachers are supposed to be the 'provider of the knowledge' and students are supposed to act as the 'passive receiver of the materials' (Aljohani, 2017). Again, this finding validates previous studies carried out in the Saudi EFL context, which indicates that teachers do not adequately appreciate the importance of providing autonomy to language learners in the classroom (Al-Awaid, 2018; Alrabai, 2016).

The qualifications held by the teachers have also been put under the lens in this study to determine their effect on creating a student-centred approach to teaching EFL in the context of Saudi Arabia. Previous studies already have an unsatisfactory view of the qualifications possessed by English teachers in the Saudi context (Al-Hazmi, 2003; Picard, 2018). However, in this study, the average response given by participant English teachers can be considered mediocre. This study indicates teachers' view regarding their own qualifications was not particularly encouraging – a finding that aligns with previous literature (Al-Zahrani & Rajab, 2017). The reason behind teachers' lack of perception and appreciation for their qualifications can be traced to their lack of self-actualisation, confidence and uncertainty about the responsibilities of their professional requirements. Recognition of one's capabilities and skills can be of uppermost importance for individuals. This enables them to accomplish their professional goals. Getting higher qualifications and having a variety of teaching skills are considered good signs of achieving self-actualisation because this will be positively reflected in their teaching techniques. For instance, qualified teachers recognise the effectiveness of applying Information Communication technology in their classrooms. In this study, a possible reason behind teachers assigning

such a low priority to this issue can be their overconfidence about ICT efficacy in education and the misconception that ICT integration is proportional to quality teaching (Figure 3, item 46-47). However, literature advocates the use of everyday materials. When considering students' motivation to learn EFL, the constructivist approach emphasises that making the concepts they learn more relevant to real life can motivate learners to be updated on their coursework (Aljohani, 2017).

The fourth attribute in this study was the teachers' perception of group activities for effective EFL teaching in a student-centred environment. It is irrefutable that group activities, tasks, assignments or even discussions can promote self-actualisation and self-determination among students from a humanist approach (Huitt, 2009). Such measures can positively impact students learning EFL (Al-Garni & Almuhammadi, 2019; Alrayah, 2018; Ochoa et al., 2016). In this study, the participant teachers were considerably more positive toward the idea of promoting group activities in their EFL teaching after considering their students' circumstances (Figure 4, items 9-10). This finding, however, is opposed to the previous literature, where the findings indicate that teachers are lukewarm to the prospect of integrating group activities into their study (Alfares, 2017).

Nevertheless, English teachers in the Saudi context were also found to be lagging behind when creating lesson plans before teaching a lesson. It was the lowest recorded attribute found in this study (Figure 4, item 1). The reason behind this has already been explained in previous literature, where it had been attributed to the lack of appropriate teacher training on the importance of strategic lesson planning teaching EFL (Alharbi, 2019). Aside from this, content overload also restricts the teachers' time management abilities, who might have been interested in drafting proper lesson plans to implement in the classroom (Alrabai, 2016).

To explore more in depth shaping teacher perspectives on promoting a student-centred approach in teaching EFL, a group of important variables were analysed. These variables included gender (male, female), location of the school (urban, rural), the status of the school (public or Government funded, private) and class size (number of students less than 30, 30-50, and more than 50). Although multiple studies have emphasised these aspects' impact in teaching EFL (Alrabai, 2016), the study's findings indicate that these factors have an insignificant influence in shaping EFL teachers' perspectives. Even though these factors were deemed statistically insignificant for the most part, it should be noted that the female participant teachers had a slightly higher positive response towards promoting these four major attributes compared to their male counterparts, especially regarding corrective behaviour and student diversity. Concerning school location and status variables, it was found that participant English teachers from rural and private schools had a slightly more positive perception of promoting the four major attributes to create a student-centred learning atmosphere. According to the class size variable, teachers with more than 30 students in their classrooms responded more positively toward group activities or small group tasks. This is not surprising, considering previous studies have also mentioned that gender, demography, status and class size impact language teachers' perception and their ability to teach EFL effectively in the Saudi context (Al-Awaid, 2018; Alrabai, 2016).

LIMITATIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

This study has used the purposeful sampling technique, which means that it is not truly representative of the actual state of things in the Kingdom of Saudi Arabia. Another limitation experienced in this study was the limited opportunity to review the responses submitted by teachers and examine them comprehensively through semi-structured interviews. The COVID-19 pandemic made it impossible to carry out face-to-face interviews for qualitative case analysis. It is recommended that further studies on this subject be undertaken to explore the specific aspects revealed by the findings of this study to rationalise the results. Additionally, help can be sought from the government's education ministry by requesting a representative sample that accurately reveals the views held by EFL teachers across the country.

CONCLUSION

The study offered a general overview of EFL teachers' perceptions of encouraging a student-centred language learning process through the lens of four major attributes that measured the behaviour exhibited by a group of teachers in the Saudi EFL context. The literature identifies student-centred learning with elements such as alternative assessment, focusing on meaning, students and teachers as co-learners, student-student interaction, curricular integration, learner autonomy, learning climate, diversity and motivation. These elements were measured under the lens of teachers' observable behaviour, clustered under four common attributes between the two concepts: assessment strategy of EFL, the communicative approach to learning EFL, teachers' qualification for teaching EFL and group activities in teaching EFL. In general, from the data, it was inevitable that teaching grammar approaches showed the tendency to be undertaken in a teacher-centred approach. However, while teachers were found to be open to the idea of a student-centred approach, some aspects like using daily objects, preparation of lesson plans, and time management had a lot of room for improvement, which could be done by reducing course overload and offering adequate and appropriate training programs. Participatory teachers' perception of their qualifications was mediocre, which meant that the

qualification standards needed improvement. To satisfy the needs of the teachers, further studies need to be undertaken on this subject.

The role played by factors like gender, location and status of the school and class size in determining the teachers' perception regarding student-centred learning was also studied. Although the results were deemed statistically insignificant for the study, it should be noted that teachers whose student count was over 30 and teachers from private and rural schools and belonging the female gender were found to be more positive towards the idea of addressing four major attributes of student-centred learning compared to their counterparts.

DECLARATION OF COMPETING INTEREST

None declared.

REFERENCES

Abahussain, M. O. (2016). Implementing communicative language teaching method in Saudi Arabia : challenges faced by formative year teachers in state schools [Unpublished Doctoral Dissertation]. University of Stirling, Scotland.

Al-Awaid, A. S. (2018). Teaching strategies in EFL environment in the secondary schools in the KSA: evaluation and remedies. Advances in Language and Literary Studies, 9(2), 50-58. http://dx.doi.org/10.7575/aiac.alls.v.9n.2p.50

AL-Garni, S. A., & Almuhammadi, A. H. (2019). The effect of using communicative language teaching activities on EFL students' speaking skills at the University of Jeddah. English Language Teaching, 12(6), 72-86. https://doi.org/10.5539/elt.v12n6p72

Al-Hazmi, S. (2003). EFL teacher preparation programs in Saudi Arabia: Trends and challenges. TESOL Quarterly, 37(2), 341-344. https://doi.org/10.2307/3588509

Al-khresheh, M. (2020a). The influence of anxiety on Saudi EFL learners' oral performance, University Aljouf humanities sciences journal, 6(8), 275-305.

Al-khresheh, M. (2020b). The impact of cultural background on listening comprehension of EFL Saudi students. Arab World English Journal, 11(3), 349- 371. https://dx.doi.org/10.24093/awej/vol11no3.22

Al-khresheh, M., & Al-Ruwaili, S. (2020c). An exploratory study on vocabulary learning strategies used by Saudi EFL learners. Journal of History Culture and Art Research, 9(2), 288-302. http://dx.doi.org/10.7596/taksad.v9i2.2616

Al-Seghayer, K. (2014). The actuality, inefficiency, and needs of EFL teacher-preparation programs in Saudi Arabia. International Journal of Applied Linguistics and English Literature, 3(1), 143–151. https://doi.org/10.7575/aiac.ijalel.v.3n.1p.143

Al-Zahrani, N.O.A., & Rajab, H. (2017). Attitudes and perceptions of Saudi EFL teachers in implementing Kingdom of Saudi Arabia's vision 2030. International Journal of English Language Education, 5(1), 83-99. https://doi.org/10.5296/ijele.v5i1.10733

Aldawood, S. F. (2016). Washback of continuous assessment in EFL classrooms in Saudi Arabia: teachers' perceptions and practices [Unpublished doctoral dissertation]. Universiti Sains.

Alfares, N. (2017). Benefits and difficulties of learning in group work in EFL classes in Saudi Arabia. English Language Teaching, 10(7), 247-256. https://doi.org/10.5539/elt.v10n7p247

Alharbi, Y. G. (2019). A review of the current status of English as a foreign language (EFL) education in Saudi Arabia. Global Journal of Education and Training, 1(2), 1-8.

Aljohani, M. (2017). Principles of "constructivism" in foreign language teaching. Journal of Literature and Art Studies, 7(1), 97-107. https://doi.org/10.17265/2159-5836/2017.01.013

Almazova, N. I., Kostina, E. A., & Khalyapina, L. P. (2016). The new position of foreign language as education for global citizenship. Novosibirsk State Pedagogical University Bulletin, 6(4), 7–17. https://doi.org/10.15293/2226-3365.1604.01

Alnujaidi, S. (2019). The difference between EFL students' preferred learning styles and EFL teachers' preferred teaching styles in Saudi Arabia. English Language Teaching, 12(1), 90-97. https://doi.org/10.5539/elt.v12n1p90

Alrabai, F. (2016). Factors underlying low achievement of Saudi EFL learners. International Journal of English Linguistics, 6(3), 21-37. https://doi.org/10.5539/ijel.v6n3p21

Alrayah, H. (2018). The effectiveness of cooperative learning activities in enhancing EFL learners' fluency. English Language Teaching, 11(4), 21-31. https://doi.org/10.5539/elt.v11n4p21

Alsalem, M. S. (2020). Language teacher cognition on English grammar assessment: Investigating EFL teachers' beliefs and practices regarding classroom-based assessment in public higher-educational institutions in Saudi Arabia [Unpublished doctoral dissertation]. University of Essex.

Assalahi, M. H. (2013). Why is the grammar-translation method still alive in the Arab world? Teachers' beliefs and its implications for EFL teacher education. Theory and Practice in Language Studies, 3(4), 589-599. https://doi.org/10.4304/tpls.3.4.589-599

Baksh, A., Mohd Sallehhudin, A. A., Tayeb, Y. A., & Norhaslinda, H. (2016). Washback effect of school-based English language assessment: A case-study on students' perceptions. Pertanika Journal of Social Sciences & Humanities, 24(3). 1087–1104.

Bell, T. R. (2005). Behaviors and attitudes of effective foreign language teachers: Results of a questionnaire study. Foreign Language Annals, 38(2), 259–270. https://doi.org/10.1111/j.1944-9720.2005.tb02490.x

Bonner, S. M., Diehl, K., & Trachtman, R. (2019). Teacher belief and agency development in bringing change to scale. Journal of Educational Change, 21(2), 363–384. https://doi.org/10.1007/s10833-019-09360-4

Creswell, J. W. (2002). Educational research: Planning, conducting, and evaluating quantitative (146-166). Prentice Hall.

Debreli, E. (2016). Pre-Service teachers' belief sources about learning and teaching: An exploration with the consideration of the educational programme nature. Higher Education Studies, 6(1), 116-127. https://doi.org/10.5539/hes.v6n1p116

Deci, E. L., & Ryan, R. M. (2010). Self-Determination. In I.B. Weiner & W.E. Craighead (Eds.), The Corsini encyclopedia of psychology. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470479216.corpsy0834

Devaki, V. (2021). Influence of behaviourist and cognitivist theories in adult language acquisition. Journal of English Language Studies, 3(1), 38-44. https://doi.org/10.31849/ELSYA.V3I1.5620

Dewaele, J. M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. Language Teaching Research, 22(6), 676–697. https://doi.org/10.1177/1362168817692161

Godwin-Jones, R. (2019). Riding the digital wilds: Learner autonomy and informal language learning. Language Learning and Technology, 23(1), 8–25. https://doi.org/10.125/44667

Huitt, W. (2009). Humanism and open education. In W. Huitt (Ed.), Educational psychology interactive. Valdosta State University.

Jacobs G.M., & Renandya W.A. (2016) Student-centred learning in ELT. In W. Renandya & H. Widodo (Eds.), English language teaching today. English Language Education (vol. 5, pp. 13-23). Springer. https://doi.org/10.1007/978-3-319-38834-2_2

Kassem, H. M. (2018). The impact of student-centered instruction on EFL learners' affect and achievement. English Language Teaching, 12(1), 134-153. https://doi.org/10.5539/elt.v12n1p134

Köksal, D., & Ulum, Ö. G. (2019). Pre-service EFL teachers' conceptions of language proficiency: Entry and exit level qualifications. Journal of Language and Linguistic Studies, 15(2), 484–495. https://doi.org/10.17263/jlls.586144

Mohammad, T., & Hazarika, Z. (2016). Difficulties of learning EFL in KSA: Writing skills in context. International Journal of English Linguistics, 6(3), 105-116. https://doi.org/10.5539/ijel.v6n3p105

Murray, G., & Lamb, T. (2018). Space, place, and autonomy in language learning: An introduction. In G.Murray & T. Lamb (Eds.), Space, place, and autonomy in language learning (pp. 1–6). Routledge.

Ochoa, C., Cabrera, P., Quiñónez, A., Castillo, L., & González, P. (2016). The effect of communicative activities on EFL learners' motivation: A case of students in the Amazon region of Ecuador. Colombian Applied Linguistics Journal, 18(2), 39. https://doi.org/10.14483/calj.v18n2.10018

Orak, D. S., & Al-khresheh, M. (2021). In between 21st century skills and constructivism in ELT: Designing a model derived from a narrative literature review. *World Journal of English Language, 11*(2), 166-176. https://doi.org/10.5430/wjel.v12n1p1

Overskeid, G. (2008). They should have thought about the consequences: The crisis of cognitivism and a second chance for behavior analysis. The Psychological Record, 58(1), 131-151. https://doi.org/10.1007/BF03395606

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Administration and policy in mental health and mental health services research, 42(5), 533-544. https://doi.org/10.1007/s10488-013-0528-y

Pathan, H., Memon, R. A., Memon, S., Khoso, A. R., & Bux, I. (2018). A Critical Review of Vygotsky's Socio-Cultural Theory in Second Language Acquisition. International Journal of English Linguistics, 8(4), 232-236. https://doi.org/10.5539/ijel.v8n4p232

Peters, G. J. Y. (2014). The alpha and the omega of scale reliability and validity. European Health Psychologist, 16(2), 56–69. https://doi.org/10.31234/osf.io/h47fv

Phillips, D. C. (2012) Behaviorism and behaviorist learning theories. In: Seel N.M. (eds) Encyclopedia of the Sciences of Learning. Springer. https://doi.org/10.1007/978-1-4419-1428-6_750

Picard, M. (2018). The future of EFL and TESOL in Saudi Arabia. In C. Moskovsky & M. Picard (Eds.), English as a foreign language in Saudi Arabia (157–177). Routledge. https://doi.org/10.4324/9781315688466-8

Pritchard, A. (2017). Ways of learning: Learning theories for the classroom. Routledge.

Rind, I. A., Mari, M. A., & Heidari-Shahreza, M. A. (2019). Analysing the impact of external examination on teaching and learning of English at the secondary level education. Cogent Education, 6(1), 1574947. https://doi.org/10.1080/2331186X.2019.1574947

Roussel, S., Joulia, D., Tricot, A., & Sweller, J. (2017). Learning subject content through a foreign language should not ignore human cognitive architecture: A cognitive load theory approach. Learning and Instruction, 52, 69-79. https://doi.org/10.1016/j.learninstruc.2017.04.007

Ryan, R. M., & Deci, E. L. (2016). Facilitating and hindering motivation, learning, and well-being in schools: Research and observations from self-determination theory. In K. R. Wentzel & D. B. Miele (Eds.), Handbook on motivation at schools (pp. 96–119). Routledge.

Schunk, D. H. (2012). Learning theories: An educational perspective (6th ed.). Pearson.

Sert, O. (2019) Classroom interaction and language teacher education. In S. Mann & S. Walsh (Eds.), The Routledge handbook of language teacher education. Routledge. https://doi.org/10.4324/9781315659824-16

Shabani, K., Khatib, M., & Ebadi, S. (2010). Vygotsky's zone of proximal development: Instructional implications and teachers' professional development. English Language Teaching, 3(4), 237-248. http://dx.doi.org/10.5539/elt.v3n4p237

Sidhu, G. K., Muthukrishnan, P., & Zhiyong, S. (2020). College English language teaching reform and EFL lecturers' teaching practice: A case study in China. Universal Journal of Educational Research, 8(1), 230–237. https://doi.org/10.13189/ujer.2020.080128

Snowman, J., McCown, R, & Biehler, R. (2012). Psychology applied to teaching (13th ed.). Wadesworth, Cengage Learning.

Song, B., & Kim, T. Y. (2017). The dynamics of demotivation and remotivation among Korean high school EFL students. System, 65, 90–103. https://doi.org/10.1016/j.system.2016.12.010

Svinicki, M. D. (2004). Learning and motivation in the postsecondary classroom. Anker Publishing Company.

Taylor, G. R., & MacKenney, L. (2008). Improving human learning in the classroom: Theories and teaching practices. R&L Education.

Tobolowsky, B. F. (2014). Paths to learning: Teaching for engagement in college. Stylus Publishing, LLC.

Tolmie, A., Muijs, D., & McAteer, E. (2011). Quantitative methods in educational and social research using SPSS. Open University Press.

United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development. In Resolution adopted by the General Assembly. United Nations General Assembly.

Watson, J. B. (2017). Behaviorism. Routledge.

Wilden, E., Porsch, R., & Schurig, M. (2020). An early start in primary EFL education and the role of teacher qualification and teaching quality. Language Teaching for Young Learners, 2(1), 28–51. https://doi.org/10.1075/ltyl.19002.wil

Wilson, B. G., & Myers, K. M. (2000). Situated cognition in theoretical and practical context. In D. H. Jonassen & S. M. Land (Eds.), Theoretical foundations of learning environments (pp. 57-88). Erlbaum.

Wozney, L., Venkatesh, V., & Abrami, P. (2006). Implementing computer technologies: Teachers' perceptions and practices. Journal of Technology and teacher education, 14(1), 173-207.

Wubante, M (2020). Review on behaviourist approach and the construction of knowledge. International Journal of English Literature and Culture, 8(6), 164-171. https//10.14662/IJELC2020.125

You, S., Kim, E., & Shin, K. (2019). Teachers' belief and efficacy toward inclusive education in early childhood settings in Korea. Sustainability, 11(5). http://dx.doi.org/10.3390/su11051489

Yi, H. (2008). The effect of class size reduction on foreign language learning: A case study. Language and Linguistics Compass, 2(6), 1089-1108. https://doi.org/10.1111/j.1749-818X.2008.00090.x

APPENDIX 1

Item Statistics

Item Statistics				
Items:	**In your opinion, the effective foreign language teacher**	**Mean**	**Std. Deviation**	**N**
1.	creates lesson plans that emphasise grammatical aspects of the TL.	2.12	.604	302
2.	teaches new complex language structures only after less complex structures have been introduced and practiced.	2.31	.709	302
3.	uses information gap activities (where students have to find out unknown information from a classmate or another source).	2.34	.752	302
4	uses small groups to help learners experience a greater degree of involvement.	2.47	.680	302
5	gives learners a time limit to complete small group activities.	2.48	.656	302
6	bases at least some part of students' grades on their actual use of the TL.	2.27	.718	302
7	bases at least some part of students' grades on completion of assigned tasks.	2.50	.598	302
8	uses student-student role play situations from the beginning of elementary language instruction.	2.53	.580	302
9	encourages students to express and discuss their needs and preferences for language learning.	2.58	.558	302
10	adjusts learning activities to meet the needs of foreign language students with a variety of interests.	2.59	.580	302
11	shows personal involvement in or enthusiasm for the TL and culture.	2.52	.625	302
12	permits learners to select their own topics for discussion.	2.52	.630	302
13	teaches foreign language students to use various learning strategies (i.e., self-evaluation, repetition, imagery, etc.).	2.55	.601	302
14	varies learning activities of foreign language instruction depending on learners' ages.	2.62	.539	302
15	uses activities and assignments that draw learners' attention to specific grammatical features.	2.60	.536	302
16	uses activities where learners need to understand a certain grammatical feature to understand the meaning of spoken or written text.	2.60	.555	302
17	simplifies his or her TL output so students can understand what is being said.	2.50	.581	302
18	thoroughly explains new grammar rules before asking students to practice the relevant structure.	2.53	.597	302
19	teaches grammar inductively (i.e., gives examples before grammatical rules).	2.61	.559	302
20	teaches appropriate hesitation or other discourse strategies to help learners gain time in conversational exchanges.	2.52	.586	302
21	exposes students to different varieties the TL.	2.37	.698	302
22	requires students to practice unfamiliar grammatical forms or patterns in substitution or transformation exercises.	2.54	.562	302
23	uses recasts (correct reformulations of students' speech) as a preferred method of corrective feedback.	2.50	.630	302
24	corrects errors as soon as possible after they occur.	2.60	.511	302
25	uses indirect cues or hints to signal errors to the learner (such as, asking them if they are sure their response is correct or using facial expressions or body language).	2.58	.558	302
26	has students act out commands or engage in other physical activity given by the teacher to practice listening comprehension in the TL.	2.47	.586	302
27	uses the TL as the predominant means of classroom communication.	2.40	.653	302
28	provides learners with concrete tasks to complete while reading or listening to texts in the TL.	2.36	.631	302
29	teaches foreign language students to use strategies to improve their vocabulary learning (e.g., memory devices or creating a mental image of the word).	2.47	.586	302
30	presents grammar rules one at a time and has student practice examples of each rule before going on to another.	2.47	.586	302
31	devotes class time to giving examples of cultural differences between target and student's native language use.	2.48	.551	302
32	teaches idiomatic expressions and language routines to help learners successfully engage in conversations in the TL.	2.45	.612	302
33	encourages learners to begin speaking in the TL only when they feel thev are ready to.	2.50	.575	302

Item Statistics				
Items:		**Mean**	**Std. Deviation**	**N**
In your opinion, the effective foreign language teacher				
34	encourages foreign language learners to speak in the TL beginning the first day of class.	2.46	.596	302
35	explains why learner responses are inaccurate when students make errors.	2.54	.550	302
36	allows students to write summaries or answer questions on reading or listening passages in English rather than the NL.	2.46	.562	302
37	grades spoken language production predominantly for grammatical accuracy.	2.52	.569	302
38	grades written language assignments predominantly for grammatical accuracy.	2.55	.531	302
39	understands the fundamentals of linguistic analysis (phonology, syntax) as they apply to the TL.	2.44	.600	302
40	uses the TL competently.	2.23	.726	302
41	provides opportunities for students to reinforce and further their knowledge of other disciplines through the foreign language.	2.48	.569	302
42	selects materials that present distinctive viewpoints that are available only through the foreign language and its cultures.	2.56	.510	302
43	provides opportunities for students to demonstrate understanding of the nature of language through comparisons of the TL and their own.	2.51	.533	302
44	provides opportunities for students to use the TL both within and beyond the school setting.	2.50	.592	302
45	teaches grammar deductively (i.e., gives grammatical rule before examples).	2.55	.531	302
46	integrates computer-aided instruction (e.g., computer-based exercises, email, the Internet, CD-ROM, etc.) into foreign language teaching.	2.57	.541	302
47	frequently uses authentic materials and realia (e.g., maps, pictures, artifacts, items of clothing, foods) to illustrate features of the TL and culture.	2.46	.525	302

https://doi.org/10.17323/jle.2022.13868

Changes of Meta-Representational Skills in Ageing: First Empirical Evidence on the Relation between Metalinguistic Competence and Attributions of Mental States

Federica Bianco [1], **Ilaria Castelli** [1], **Carmen Belacchi** [2]

[1] University of Bergamo, Bergamo, Italy

[2] University of Urbino Carlo Bo, Urbino, Italy

ABSTRACT

Background. The present paper focuses on meta-representational changes occurring in ageing by studying the decline in Definitional Competence, an ability so far little studied in this period of life.

Purpose. The paper hypothesises a relation between Theory of Mind (ToM) and Definitional Competence, in a view that posits the former as a preparatory and facilitating competence for a more complex linguistic production, that is lexicographic definition. The effects of levels of education on the decline in ageing for Definitional Competence and ToM are also investigated.

Methods. We recruited 24 adults (age range 21–55), 25 young-old adults (age range 60–70) and 25 old-old adults (age range 71–85) and administered them the Eyes task to measure ToM and the Co.De. Scale to assess Definitional Competence.

Results. Results suggest that Definitional Competence declines earlier on than ToM, mirroring the well-known process according to which in semantic knowledge, during ageing, taxonomic relations are lost before thematic ones. Our results also show that better levels of education are associated with better performance in both our key constructs and that ToM predicts Definitional Competence, in line with our expectations.

Conclusion. The paper offers one of the first systematic studies on the changes in Definitional Competence during the last phase of life and it provides theoretical insights into the relation between ToM and Definitional Competence in ageing.

Implications. The paper is informative for future interventions aimed at enhancing linguistic and metalinguistic skills in ageing through the preservation of better levels of awareness and the assumption of a decentralised perspective in interpersonal communication.

KEYWORDS

definitional competence, theory of mind, meta-representation, ageing, levels of education

Citation: Bianco, F., Castelli, I., & Belacchi, C. (2022). Changes of Meta-Representational Skills in Ageing: First Empirical Evidence on the Relation between Metalinguistic Competence and Attributions of Mental States. *Journal of Language and Education, 8*(3), 40-51. https://doi.org/10.17323/jle.2022.13868

Correspondence:
Federica Bianco
federica.bianco@unibg.it

Received: February 12, 2022
Accepted: September 07, 2022
Published: September 30, 2022

INTRODUCTION

Meta-representation ability, the skill that permits one to implicitly and explicitly grasp and share one's own and other internal states, is a distinctive feature of the human species and finds its highest manifestation in language (Tomasello, Carpenter, Call, Behne, & Moll, 2005). Cognitively speaking, language is a powerful means of representation that organises world knowledge into permanent and systematic mental structures and enables the sharing of this knowledge among all members of a given culture (Tomasello, 1999). Language also facilitates continual analytical reflection on experience and allows different hypothetical scenarios to be represented. From a communication point of view, language permits the negotiation of meanings, plans and solutions, as well as the explanation of individual mental states. In other ways, language and, crucially, the ability to define words allows us to

make explicit the implicit representations (both personal and cultural) according to the Representational Rewriting Model by Karmiloff-Smith (1992).

The prerequisite par excellence of language development consists in the ability to recognise that other individuals have minds that function in a similar way to ours (Tomasello, 2001) and in understanding other's intentions to pursue the "same" means for the "same" end (Tomasello, 2003, p.33). Other skills that seem particularly important for language acquisition involve various kinds of broadly defined pattern-finding categorisation (Rosh et al., 1976).

THEORETICAL BACKGROUND

The ability to understand others' internal representation behind manifest behaviour is called Theory of Mind (ToM) (Premack & Woodruff, 1978; Wimmer & Perner, 1983). Language and ToM share high-level cognitive and meta-representational components, and this common substrate is thought to explain how the acquisition of each of these competencies is closely linked to the acquisition of the other (Astington & Baird, 2005; Tomasello, 2001). In particular, among linguistic skills, the Definitional Competence, the ability to make verbally explicit the meaning of a given word (*definiendum*) by means of other words (*definiens*), implies metalinguistic and meta-representational capacities (Benelli et al., 2006), also crucial for ToM competence. Indeed, producing good lexicographic definitions demands not only the metalinguistic capacity to follow conventional rules for making word meanings explicit, the selection of relevant semantic content components, and the use of appropriate linguistic forms, but also the ability to take simultaneously into account one's own and others' information needs and goals. Despite the fact that the two skills studied here (namely *'ToM' and 'Definitional Competence'*) seem to share the same underlying meta-representative process, they are also deeply different in their nature and thus can be studied as different constructs. ToM competence usually acts at an implicit (unless the subject is asked to verbally explicit the contents of their inferences upon mental states, such as in the second- and third-order false belief tasks or in some conversations) and contextualised level, while the Definitional Competence requires the mastering of explicit, de-contextualised and intersubjective forms of sharing mental representations. At a theoretical level, in order to provide the interlocutor with adequate definitions through a conventional and effective verbal exchange, the speaker needs to put themselves in the shoes of the interlocutor, assuming their cognitive and affective representation of the world (the core of ToM). However, few studies have investigated the meta-representational components of Definitional Competence (e.g. Belacchi & Benelli, 2021), requiring awareness of both the conventional value of definitions and shared meanings and no less of the interlocutors' informational needs and, more generally, of the ability to assume other perspectives.

From the literature in this domain, it is clear that, even if many previous empirical works examined the reciprocal links between language and ToM acquisition (see, for example, Astington & Baird, 2005; Tomasello, 2001), we still have little understanding of how the ability to define words and ToM are related in different phases of life. We already know that these kinds of meta-representational abilities are influenced by age, cognitive abilities and level of education (Benelli et al., 2006; Gini, Benelli and Belacchi, 2004; Li et al., 2013; Pardini e Nichelli 2009; Phillips et al., 2002; Slessor et al., 2007), but the relation between implicit (automatic and unconscious) and explicit (deliberate and conscious) levels/mechanisms that characterise the structure, functioning and specific manifestations of different meta-representational skills is still an open question. In particular, following both previous theoretical approaches (e.g., Malle, 2022) and preliminary evidence (Belacchi 2004; Belacchi, 2019; Roselli, Valentini, & Belacchi, 2019) on this issue, we are interested in testing whether the ability to implicitly represent one's own and others' psychic states (i.e., ToM as measured by tasks assessing automatic attribution of inner states) favours the mastery of language in its most advanced and explicit forms, as in the case of Definitional Competence.

In the present paper we aim to provide one of the first contributions in this domain by focusing on aging. This period of the life-span can be particularly informative on the nature of the relation between ToM and Definitional Competence, as we know from a rich body of previous research that ToM undergoes deep transformations during the last phase of human life (Henry, Louis, Phillips, Ruffman, & Bailey, 2012; Hughes et al., 2019) and that crystalised knowledge tends to be maintained while fluid intelligence seems to decline in this period of life. On the other hand, recent literature on ageing has elucidated that some aspects of language skills decline across ageing while others are mainly preserved. In particular, it seems that productive aspects of language are more vulnerable to the effect of age (Burke & Shafto, 2008; Zhang, Eppes, & Diaz, 2019). Moreover, previous research suggests that the thematic memory organisation is better preserved in older adults than the taxonomic organisation (Belacchi & Artuso, 2018; Belacchi, Artuso, & Palladino, 2022), so it is plausible that those parts of language skills that use taxonomic categorisation are particularly impaired in the ageing process. Moreover, Artuso & Belacchi (2021) showed that the elderly seem to use thematic knowledge to support expository text comprehension and not the taxonomic one, which would work better in the light of similar common processes (such as abstract logical relationships and de-contextualisation). This confirms that hierarchically organised semantic structures (i.e., taxonomic ones) are not specifically useful in comprehending complex texts for older adults, conversely to what has been observed in adults.

In the following paragraph, we will better analyse each of our target skills and provide a framework to understand the main developmental changes from a life span perspective.

Definitional Competence: Its Nature and Developmental Trend

As stated above, language is a powerful means of representation that codes the most relevant knowledge about the world, allowing to share this knowledge among all members of a given culture. From a communicative point of view, language allows negotiation of meanings, plans and solutions, as well as an explanation of individual mental states such as notions, beliefs, emotions and needs (Nelson, 1996; Tomasello, 2001).

One of the most powerful ways to understand the most abstract, meta-linguistic and meta-representational aspects of language is to analyse the ability to define words: definitions require a well-established, conventionally shared semantic knowledge and linguistic expression rules, being part of the intersubjective "common ground" (Tomasello, 2008). The prototypical format of a lexicographic definition is the Aristotelian formula: "An X is a Y that Z", which involves lexical, semantic and syntactic abilities to express the meanings of words (Belacchi & Benelli, 2007; Belacchi & Benelli, 2017; Benelli, Belacchi, Gini & Lucangeli, 2006). Traditional literature on conceptual and definitional development has shown that different definitional levels exist according to the kinds of answers children provide (Litowitz, 1977; Snow, 1990; Benelli & Belacchi, 2017). According to a developmental trend, prototypical definitional answers range from the so-called HAS structure ("a chair *has* four legs"), which simply describes the sensorial and experiential aspects of referents, to the so-called ISA structure ("a chair *is a* piece of furniture"), with a taxonomic organisation of conceptual knowledge. This syntagmatic to paradigmatic shift occurs around six to seven years of age.

Lexicographic definitions are an expression of explicit metalinguistic competence: to define implies not only knowing *what to say* but also *how to say* it (McGee-Bidlack, 1991; Snow, 1990; Watson, 1995; Benelli et al., 2006). Producing good definitions also has meta-communicative and meta-representational components; it requires awareness of the conventional value of definitions and shared meanings and interlocutors' informational needs.

Recently, a scale was developed (Benelli, et al. 2006; Belacchi & Benelli, 2017; Belacchi & Benelli, 2021) to assess Definitional Competence (Co.De. Scale). This scale identifies definitional levels, emphasising the formal aspect of the answers and analytically expanding the semantic content that is not directly evaluable. Specifically, the higher the scale levels, the better they represent the requirements of high-quality definitions. High-quality definitions (that are successful in making a reciprocal correspondence between *definiendum* e *definiens* at the semantic content level) have to respect the following formal rules: illustration of a meaning of a word by using other words (no-tautology rule); provision of more words than the stimulus item (phrasal extension rule); provision of a verbal expression to make explicit the semantic equivalence (copula rule); provision of a syntactically autonomous and correct linguistic structure (syntactic autonomy rule) (Belacchi & Benelli, 2021). See Table 1 for more details on the levels of the scale.

Previous research (e.g. Belacchi & Benelli, 2007; 2021; Benelli & Belacchi, 2017) has shown an increase in both formal correctness and completeness of the linguistic structures used to define words from preschool age to adulthood. The same age-related trends were found for nouns, verbs and adjectives, and concrete and abstract nouns. It was also demonstrated that there is a high concurrent validity of the Definitional Scale with traditional measures of intelligence, both in typical and atypical populations. In particular, mental and chronological age are predictors of Definitional Competence in children with intellectual disabilities (Belacchi & Benelli, 1999; 2007). Moreover, definitional skills, as a measure of de-contextualised language abilities, are closely related to school success and literacy (Artuso, Palladino, Valentini & Belacchi, 2022; Gini, Benelli & Belacchi, 2004; Gutierrez-Clellen & DeCurtis, 1999; Sinclair, 1986; Snow, 1990).

Table 1

Definitional Levels, Prototypical Answers and Scores for the Definition of the Word "Donkey"

Levels	Types of answers	Score
0. Non-definition	Non-verbal answers	0
I. Pre-definition	One-word answers, mostly associations (e.g. *donkey-> ears*)	1
II. Quasi-definition	Initial formulation of sentences without autonomous forms (e.g. *donkey-> with the long ears; when it brays*)	2
III. Narrative/descriptive definition	Formally correct and autonomous sentences with narrative/descriptive content (e.g. *donkey brays; donkey is mild*)	3
IV. Categorical definition	Formally correct and autonomous sentences in simply categorical/synonymic form (e.g. *The donkey is an animal*)	4
V Partial Aristotelian definition	Formal correctness without semantic equivalence (e.g. *The moon is a planet in the solar system*)	5
VI. Aristotelian, metalinguistic definition	Formal and semantic correctness and equivalence (e.g. *A donkey is an animal that brays*)	6

To our knowledge, this is the first complete study on Definitional Competence in ageing, although some preliminary results have also been presented at conferences (Belacchi, 2019).

Theory of Mind: Its Nature and Developmental Trend

In ToM reasoning, being able means being able to acknowledge the existence of mental states and to explain social actions in terms of inner states guiding the behaviour (Premack & Woodrooff, 1978; Wimmer & Perner, 1983). The more a person is competent in ToM, the more they are able to correctly ascribe a particular mental state (i.e., intention, emotion, desire, thought) to a focus agent, and the more they tend to adopt a mental-state lexicon to describe other people and their actions (Hughes, 2011). Studies on individual differences in ToM consistently show that this ability is fundamental for having a satisfying social life across an entire life span (Lecce et al., 2017; Slaughter, Imuta, Peterson, & Henry, 2015), and recent research has also demonstrated that ToM plays a role in the cognitive adjustment of the subject (Miller, 2012; Wellman, 2018).

The heart of ToM competence is achieved around four years of age when the child becomes capable of reasoning on how a person may be guided in their actions by wrong thoughts on reality (Wellman, 2018). Then, the competence becomes more and more mature across middle childhood (Bianco, Lombardi, et al., 2019; Bianco et al., 2021; Lecce, Bianco, Devine, & Hughes, 2017), preadolescence (Devine & Hughes, 2013) and adolescence (Valle, Massaro, Castelli, & Marchetti, 2015). Interestingly, a meta-analysis of ToM studies in ageing showed that older adults score significantly poorer than their younger counterparts across a wide range of measures varying on the modality of assessment, with a mean effect size of the age-effect moderate in magnitude (Henry, Phillips, Ru□man, & Bailey, 2012). This ToM impairment begins at about 60 years of age and becomes more accentuated after 80 years of age (Charlton, Barrick, Markus, & Morris, 2009). Sometimes, the decline was observed at the level of brain activation (e.g., Castelli et al., 2010).

Notwithstanding these acknowledgements, our comprehension of how ToM changes during ageing is far from complete. In particular, we do not know how the decline observed after 60 years of age is bound to the functioning of other near cognitive domains and if sex differences may occur (the role of sex in the ability to reason on inner states remains an open question, for example, see Baron Cohen, 2001 vs Serafin & Surian, 2004).

Literature also showed an association between years of formal education and ToM competence (Li, Wang, Wang, Tao, Xie, & Cheng, 2013; Pezzuti, Longobardi, Milletti, & Ovidi, 2011).

Research Questions and Hypotheses

The present contribution aims to understand better the relationship between Definitional Competence and ToM ability in ageing, but also to add evidence on how these competencies unfold over time during the last phase of life and how they vary depending on levels of education. We expect to find an effect of age and education on the two competencies, in line with the literature reported above. Specifically, we hypothesise a decrease in the two sets of abilities related to age and an increase in the two sets of skills related to more years of formal education. As literature reports mix findings for the presence of sex differences in meta-representational skills (Astington & Jenkins, 1995; Baron-Cohen et al., 1999; Bosacki & Astington, 1999; O'Hare, Bremner, Nash, Happé, & Pettigrew, 2009), we checked for this potential effect without an a-priori hypothesis. Our final expectations concern the nature of the relationship between Definitional Competence and ToM. As our ToM measure mainly implies implicit representations, while Definitional Competence always involves explicit representations, we expected the former to act as a predictor of the latter, beyond the influence of age and schooling. To be more stringent, in analysing the relation between ToM and Definitional Competence, we also considered the type of words to be defined (words that do refer to vs words that do not refer to inner states).

METHODS

Participants

Seventy-four subjects (37 males) were recruited by local advertising (e.g., in recreational centres) in a region in the centre of Italy. Informed written consent for participation in the research was collected. All participants were treated in accordance with the ethical guidelines for research provided by the Declaration of Helsinki (World Medical Association, 2014), the American Psychological Association (APA, 2020) and the Italian Psychological Association (AIP, 2015).

None of them reported severe health problems of any kind (i.e. no anxiety, depression, sleep disturbances or other medical conditions). All participants were Italian native speakers and scored above the 24-point cut-off of the Mini-Mental State Examination (MMSE) (Folstein, Folstein, &

[1] APA American Psychological Association (2020). Publication Manual 7th Edition. American Psychological Association.

[2] AIP Italian Psychological Association (2015). Codice etico [Ethics code]. https://www.aipass.org/sites/default/files/codice%20etico%20AIP%20rev_.pdf

McHugh, 1975). The sample was composed of 24 adults (age range 21–55, *M* = 38.29, *SD* = 11.53), 25 young-old adults (age range 60–70, *M* = 65.16, *SD* = 2.97) and 25 old-old adults (age range 71–85, *M* = 76.80, *SD* = 3.40). In the group of adults, 18 out of 24 participants reported nine or more years of formal education; in the group of young-old adults, 11 out of 25 reported nine or more years of formal education; in the old-old group, only two participants reported nine or more years of formal education. All participants were unpaid volunteers.

Measures

Theory of Mind

To index ToM, we administered the Eyes Test Adult Version (Baron-Cohen et al., 1997, 2001). This test assesses the ability to attribute complex emotional and epistemic mental states from the eyes. For each item, participants are required to select which of four words best describes the thoughts or feelings expressed in a picture depicting a person's eye region. The test is made up of 36 items. For each correct answer, a score of 1 is given. The cut-off score is 13. The total score ranges from 0 to 36.

Definitional Competence

To assess Definitional Competence (Def), we used the Co.De. Scale (Belacchi & Benelli, 2021) articulated into seven definitional levels (while the previous one was articulated into five levels). This new version allows a more analytical assessment of Definitional Competence (see Table 1). The subject is presented with 32 target words (8 nouns, 8 verbs, 8 adjectives and 8 emotion terms) varying in terms of concreteness/abstraction and in being/not being words referring to mental states (ToM words). Items were presented in random order. For each item, participants are asked: "What does (stimulus word) mean?". Each definition, orally provided by participants, was rated on a 7-level scale (score range: 0-6). According to the scale, the more correct, syntactically articulated and organised a participant's definition was, the higher the definitional level attributed (see Table 1 for information about the scoring used). Given that the frequency in the use of nouns/verbs/adjectives that refer to one's own or others' mental states is an indicator of ToM competence (De Rosnay & Hughes, 2006), we separated words that refer to inner states (i.e., thoughts, desires, emotions) from other words. This allowed us to investigate if the hypothesised link between Definitional Competence and ToM is general or specific for words referring to mental states. We considered the following Italian items as ToM words: orgoglio, tristezza, rabbia, vergogna, invidia, senso di colpa, gioia, paura, premura, innocente, pensare, spionaggio, tollerare, costringere, rivalità, frustrare (in English: pride, sadness, anger, shame, envy, guilt, joy, fear, caring, innocent, to think, spying, to tolerate, to force, rivalry, to frustrate). We considered the following Italian items as non-ToM words: asino, battere, scottare, pagliaccio, magro, rotondo, rischioso. unire, arancia, educato, emigrare, biondo, abilità, liscio, ombrello, contagioso (in English: donkey, to beat, to burn, clown, thin, round, risky, to join, orange, polite, to emigrate, blond, skill, smooth, umbrella, contagious). Participants received mean definitional scores (range 0-6) for total, non-ToM, and ToM words.

Analysis

Statistical analyses were performed with SPSS version 19. We started by running preliminary analyses on the distribution of scores on the Co.De. Scale; we then checked if sex affected our key measures in order to know if including sex (as control variable) in the following analysis or excluding it. We conducted univariate ANOVAs to examine effects of age and years of education on both Definitional Competence and theory of mind. Pairwise comparisons, corrected for Bonferroni indices, were performed in order to break down significant effects. To investigate the possible relations between definitional skills and ToM, we first conducted correlational analyses. Then, to better understand the impact of ToM on Definitional Competence, we run a hierarchical regression analysis entering, at Step 1, age and years of education, and, at Step 2, ToM scores. To be more stringent, we conducted analyses considering firstly all the items of the Co.De Scale and then separating words concerning mental states (ToM words) and words not concerning mental states (non-ToM words).

As our sample was not balanced concerning years of education, we could not test for age X education interaction.

Procedure

At the first stage we collected Informed written consent for participation in the research. At the beginning of the study subjects were administered the Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) and asked for information about years of education and potential medical issues.

The focus tasks were administered in a fixed order: the Eyes task, then the Co.De Scale. The evaluation took place during a single session lasting about one hour at the homes of participants.

RESULTS

The results section is organised into four sections: a section focused on preliminary analyses to provide information on the distribution of Definitional Competence scores across

[3] The data that support the findings of this study are available from the corresponding author, [FB], upon reasonable request.

age-groups and to check the potential effect of sex on our key measures via ANOVA; a section devoted to understand the effect of age and education on Definitional Competence via ANOVAs; a section devoted to understand the effect of age and education on theory of mind with ANOVAs; a final section where assessing relations between performance on Co.De. Scale and performance on ToM task through correlational and regression analysis.

Preliminary Analyses

From a preliminary analysis of non-verbal answers (Level 0 of the Co.De. Scale), we found 0.4% of answers of this type in the adult group, 1.13% in the young-old group, and 3.13% in the old-old group. Figures 1 and 2 show, for each age group, the proportions and distribution of responses for each verbal level of the Co.De. Scale.

Analyses showed that neither did sex have an effect on ToM, $F = 0.29$, $p = .60$, nor on Definitional Competence (total words/ToM words/non-ToM words), $0.002 < Fs < 0.42$, $.52 < ps < .96$. For this reason, we did not have a control for sex in the next analyses.

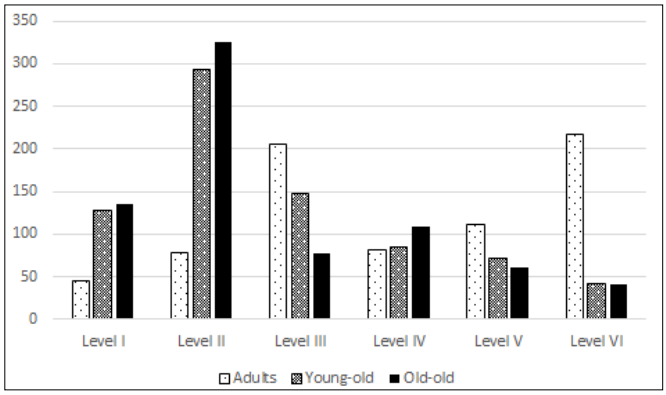
Differences in Definitional Competence scores as a function of age and education

Please note that the following analyses were conducted only on verbal answers (from Level 1 of the Co.De. Scale). Descriptive statistics for Definitional Competence in each age group and each level of education are shown in Table 3.

Analyses revealed the presence of age differences, $F(2, .71) = 24.03$, $p < .001$, partial $\eta^2 = .40$, with the group of adults outperforming both the young-old adult one, $p < .001$, 95% CI [0.76, 1.89], and the old-old one, $p < .001$, 95% CI [0.89, 2.02]. No differences emerged between the two groups of aged people, p = 1.0. The same age group effect was retrieved when we analysed effects on non-ToM words only, $F(2, 71) = 14.58$, $p < .001$, partial $\eta^2 = .29$, or on ToM-words only, $F(2, 71) = 28.51$, $p < .001$, partial $\eta^2 = .45$. Analyses showed an effect of years of education, $F(1, 72) = 29.03$, $p < .001$, partial $\eta^2 = .29$. Pairwise comparisons with the Bonferroni method also showed that people with "higher levels of education" (calculated as more than 8 years of education) outperformed people with "lower levels of education" on Definitional Competence, $p < .001$, 95% CI [0.70, 1.53].

Figure 1

Frequency Distribution of Verbal Answers Across the Levels of the Co.De. Scale as a Function of Age



Differences in ToM scores as a function of age and education

Descriptive statistics for ToM scores in each age group and each level of education are shown in Table 3. Analyses revealed the presence of age differences, $F(2, 71) = 6.05$, $p = .004$, partial $\eta^2 = .15$. Pairwise comparisons, corrected for Bonferroni indices, showed significant differences only between the adult and old-old groups, $p = .003$, 95% CI [0.46, 0.27].

In a further ANOVA, the main effect of education was significant, $F(1, 72) = 16.62$, $p < .001$, partial $\eta^2 = .19$, with more formally educated people performing better than those less educated as shown by pairwise comparisons adopting the Bonferroni method, $p < .001$, 95% CI [0.08, 0.22].

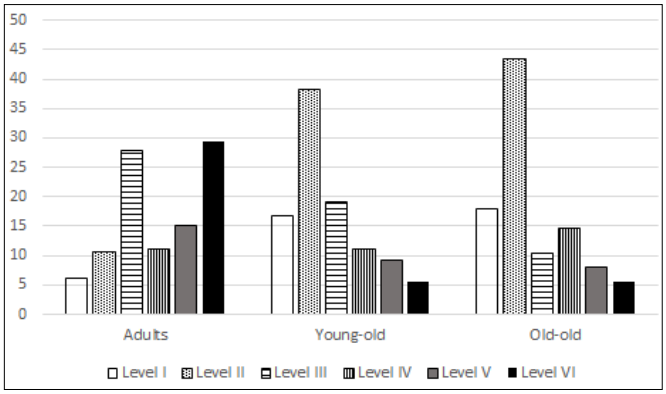
Relationship between performance on Co.De. Scale and performance on ToM task

Please note that the following analyses were conducted only on verbal answers (from Level 1 of the Co.De. Scale). We found a bivariate correlation (see Table 4) between the two skills, $r(74) = .51$, $p < .001$. Crucially, this relation is still significant, when we consider only non-ToM words, $r(74) = .46$, $p < .001$

In our first regression analysis on the Co.De Scale scores, at Step 1, we entered age and years of education, at Step 2, ToM scores. Step 1 was significant, $F(2, .71) = 51.61$, $p < .001$. Age resulted as a significant predictor, $\beta = -0.52$, $t = -4.65$, $p < .001$, and also the level of education, $\beta = 0.30$, $t = 2.68$, $p = .009$. Interestingly, Step 2 led to a significant increase

Figure 2

Percentage Distribution of Verbal Answers Across the Levels of the Co.De. Scale in Each Age Group



in the variance explained, $\Delta F(1, 70) = 5.15$, $p = .026$, with ToM uniquely predicting 2.8% of the variance in Definitional Competence scores. When we investigated the impact of ToM on Definitional Competence, separating non-ToM and ToM words, the pattern of results was similar, even if for non-ToM words, the effect of ToM was marginally significant (please see Table 5, Table 6, and Table 7 for further details on regression analyses).

Table 3

Descriptive Statistics on Definitional Competence and ToM Scores Across Age Groups and Levels of Education

	Adult Group M (SD)	Young-old Group M (SD)	Old-old Group M (SD)	High Level of Education M (SD)	Low Level of Education M (SD)
Definitional Competence (total words)	4.07 (1.09)	2.74 (0.63)	2.62 (0.64)	3.78 (1.20)	2.66 (0.54)
Definitional Competence (non-ToM words)	4.05 (1.02)	3.01 (0.70)	2.97 (0.59)	3.79 /1.14)	3.0 (0.54)
Definitional Competence (ToM words)	4.09 (1.25)	2.47 (0.64)	2.26 (0.77)	3.76 (1.34)	2.32 (0.65)
ToM	0.61 (0.16)	0.53 (0.16)	0.46 (0.15)	0.62 (0.14)	0.47 (0.16)

Note: We considered nine or more years of education as a "high level of education" and all other values as a "low level of education"

Table 4

Correlations Among Focus and Control Variables

	Age	Education	ToM	Def.Comp. Total words	Def.Comp. non-TOM	Def. Comp. ToM
Age	1	-.74***	-.42***	-.74***	-.66***	-.75***
Education		1	.49***	.68***	.60***	.71***
ToM			1	.51***	46***	.52***
Def.Comp. - Total words				1	.95***	.97***
Def.Comp. – non-TOM					1	.85***
Def. Comp. – ToM						1

*Note: + p < .10, * p < .05, ** p < .01, *** p < .001*

Table 5

Hierarchical Regression Predicting Definitional Competence (Total words)

		B	SE	β	p
Step 1	Age	-.03	.007	-.52	<.001
	Education	.07	.03	.30	.009
Step 2	Age	-.03	.006	-.50	<.001
	Education	.05	.03	.22	.054
	ToM	1.18	.52	.19	.026

R^2 = .59 for Step 1; ΔR2 = .03 for Step 2

Table 6

Hierarchical Regression Predicting Definitional Competence (non-ToM words)

		B	SE	β	p
Step 1	Age	-.03	.007	-.50	<.001
	Education	.05	.03	.22	.086
Step 2	Age	-.03	.007	-.48	<.001
	Education	.03	.03	.15	.262
	ToM	1.03	.54	.19	.062

R^2 = .46 for Step 1; ΔR2 = .03 for Step 2

Table 7

Hierarchical Regression Predicting Definitional Competence (ToM words)

		B	SE	β	p
Step 1	Age	-.04	.008	-.51	<.001
	Education	.10	.03	.34	.003
Step 2	Age	-.03	.007	-.48	<.001
	Education	.08	.03	.27	.019
	ToM	1.34	.59	.19	.027

R^2 = .62 for Step 1; ΔR2 = .03 for Step 2

DISCUSSION

The present study aimed to analyse for the first time in literature the relationship between two meta-representational skills, namely Definitional Competence and ToM, in ageing and to explore the possible analogous or different trends of developmental changes in different age groups.

Our results on the Co.De. Scale for our youngest group are in line with the mean normative scores obtained by adult people in the standardisation study conducted on 2262 Italian subjects. However, we also found a decline associated with ageing, which is worth further discussion. Even if literature consistently claims that crystallised knowledge is basically maintained across ageing (Bowles & Salthouse, 2008; Burke & Shafto, 2008; Wang & Kaufman, 1993), our results suggest that specific aspects of language, such as Definitional Competence, may be impaired. This pattern of findings can be explained considering that Definitional Competence requires a taxonomic organisation of representations, which tends to decline in the elderly, compared with better maintenance of thematic organisation (Belacchi & Artuso, 2018). A similar decline was also detected for ToM, although less marked and delayed in the age groups considered.

Our first proposal of explanation is that the decline in Definitional Competence could be due to the skills that underlie a competent use of such knowledge. Our work goes in this direction, confirming a relationship between Definitional Competence and ToM, and we already know from a large body of previous research that ToM seems to decline after 60 years of age (for a review, see Henry et al., 2013). A second possibility, not addressed here, is that the observed impairment is due to the well-known scarce executive functioning and the changes in brain connectivity (Cabinio et al., 2015.; Hughes et al., 2019) occurring in ageing. A third explanation concerns the possibility that the concurrent decline in ToM and definitional skills may reflect a more general decline in the shared underlying meta-representational domain, and indeed, we retrieved a significant association between performance on both the ToM and Definitional Competence tests, though they required quite different language skills: the Eyes Test simply requires activation of the correct name of the emotion observed, while the Co.De. Scale demands to recover and make analytically explicit the components of meaning of the terms through complex and more de-contextualised linguistic structures.

Finally, the decline in Definitional Competence might mirror a more generable vulnerability in language production during ageing. This vulnerability has been previously detected for word retrieval and off-topic speech (Burke & Shafto, 2008; Zhang, Eppes, & Diaz, 2019). However, we should note that our work reports a significant difference between the adult group and the group of aged people that for Definitional Competence emerges from the beginning of ageing, while for ToM is evident only for the group of old-old people. If we look at the developmental literature, there is consensus that basic ToM skills are mastered before Definitional Competence, which is a more complex and specialised ability. Indeed, if the core of ToM competence is achieved around three to four years of age (Wellman & Liu, 2004), it is not until seven years of age that children are able to provide canonical definitions by superordinating categorical terms (Benelli & Belacchi, 2017). This symmetric temporally reciprocal loss could be a phenomenon similar to what previous studies have observed in the case of the deterioration of semantic knowledge, in which taxonomic relations (that are acquired later on) are lost before thematic ones (Rogers & McClelland, 2004). An intriguing explanation that can make sense of this evidence is the possibility that Definitional Competence may decline earlier, as it is a more sovraordinate ability, while ToM, a more basic skill, declines later, as it is crucially linked to survival needs.

To resume, while sex does not affect the meta-representative abilities investigated here, in regards to the effect of age, it can be said that the ability to recognise emotions by reading gaze is relatively preserved in ageing, compared to the ability to define words in general, and emotions in particular. With regards to the effect of education, higher levels of education are associated with better performance on both ToM and Definitional Competence, suggesting that education can preserve a decline in different areas of cognitive processing.

We found that performance on the Eyes Test predicted performance on the Co.De. Scale over and above that is explained by education and age (please note, as a limit of our work, that these last factors explained a bigger variance than ToM). This was in line with our expectations, based on the assumption that the ability to name and recognise internal states is a preparatory and facilitating competence for a more complex meta-linguistic and meta-representational word production, i.e. lexicographic definition. The latter requires, indeed, the evocation and phrasal articulation of the meanings of words and, specifically, of terms that designate inner states (in the parts of the task that refer to ToM words). An explanation for this result is that the Eyes task measures a meta-representative ability of an implicit type, while the Co.De. Scale requires a capacity of an explicit, more evolved type, according to the representational rewriting model of Karmiloff-Smith (1992). This model considers psychological development in different domains of knowledge as a progressive, continuous transformation of the information already stored (by innate or acquired ways), recoding them in different formats that are gradually more complex, articulated and conventional – in other words, more and more explicit. On the basis of these findings, it seems that the better one's ToM capacity, the better one correctly attributes mental states and uses an adequate emotional lexicon to describe people and their actions (Hughes, 2001), con-

firming and strengthening the strong association found in previous studies between performance on ToM and simple vocabulary measures (Peterson & Miller, 2012). In the light of our results, we can add that the perspectives of others, as measured by the ToM task, enable us to define mental representations, such as the meanings of words. Therefore, our results confirm that in ageing, the implicit understanding of the intentions of others can be a determining requirement for maintaining Definitional Competence as it is, reciprocally, for learning words in the early stages of development (Baldwin, 1993; Tomasello & Farrar, 1986).

There are some limitations in our work which warrant further mentioning, as they imply guidance for further research. The first limit was the lack of other ToM measures, beyond the Eyes test that requires subjects to give a verbal answer, i.e. naming an inner state. Future studies should adopt various measures of ToM tasks employing different modalities of answers to provide more reliable estimates of the relation between the two constructs, ruling out the possibility that we are observing an artifactual association due to the specific choice of instruments. The second one regards the design of our work and the conclusions we can formulate based on this. Our work is indeed cross-sectional, and even if we found that ToM performance predicts variance in Definitional Competence scores, future longitudinal designs or experimental manipulations in a training design are required to confirm the pattern of influences. A final limitation of the current work regards the lack of balance in our groups for years of education and, therefore, the impossibility to fully shed light on the interplay among age, education and ToM in determining levels of Definitional Competence in ageing. Future research should better consider this aspect and adopt a more stringent approach to this issue.

Given the importance of language in social relations (Agha, 2006), future research should study how the decline we observed in Definitional Competence may impact the quality of life of older adults, especially regarding social exchanges. In this respect, it is worth mentioning that previous studies showed that it is possible to intervene in order to help older adults maintain their competence in socio-cognitive skills (Lecce, Bottiroli, Bianco, Rosi, & Cavallini, 2015; Rosi, Bottiroli, Cavallini, Bianco, & Lecce, 2016; Rossetto et al., 2020) and in pragmatics (Bambini et al., 2020).

CONCLUSION

To conclude, the current paper offers one of the first systematic studies on the changes in Definitional Competence during the last phase of life and provides theoretical insights into the relation between ToM and Definitional Competence. Specifically, these results suggest that Definitional Competence declines earlier than ToM, that better levels of education are associated with better performance in both our key constructs and that ToM predicts Definitional Competence, in line with our view that posits ToM as a preparatory skill for lexicographic definition. In an applicative view, our work is informative for future interventions aimed at enhancing linguistic and metalinguistic skills in ageing through the preservation of better levels of awareness and the assumption of a decentralised perspective in interpersonal communication.

DECLARATION OF COMPETING INTEREST

None declared.

REFERENCES

Agha, A. (2006). *Language and social relations. Cambridge University Press.*

Artuso, C., & Belacchi, C. (2021). Semantic memory and reading comprehension: the relationship through adulthood and aging. Aging Clinical and Experimental Research, 33(8), 2261-2271. https://doi.org/10.1007/s40520-020-01771-1

Artuso, C., Palladino, P., Valentini, P. & Belacchi C. (2022). Definitional skills as a bridge towards school achievement. *Sustainability*, *14,* 286. https://doi.org/10.3390/su14010286

Astington, J. W., & Baird, J. A. (Eds.). (2005). Why language matters for theory of mind. Oxford University Press.

Astington, J. W., & Jenkins, J. M. (1995). *Theory of mind development and social understanding. Cognition and Emotion, 9*, 151-165. https://doi.org/10.1080/02699939508409006

Baldwin D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 20*, 395–418. https://doi.org/10.1017/S0305000900008345

Bambini, V., Tonini, E., Ceccato, I., Lecce, S., Marocchini, E., & Cavallini, E. (2020). How to improve social communication in aging: Pragmatic and cognitive interventions. *Brain and Language, 211, 104864.* https://doi.org/10.1016/j.bandl.2020.104864

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry, 38, 813–822.* https://doi.org/10.1111/j.1469-7610.1997.tb01599.x

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children with Asperger syndrome or high-functioning autism. *Journal of Autism and Development Disorders, 29, 407-418.* https://doi.org/10.1023/a:1023035012436

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child and Psychology Psychiatry, 42, 241–251.* https://doi.org/10.1111/1469-7610.00715

Belacchi C. (2004, September 20-23). *Sviluppo della competenza definitoria e teoria della mente* [Development of definitional competence and theory of mind]. XVIII Congresso Nazionale di Psicologia AIP, Sezione di Psicologia dello Sviluppo, Sciacca, Italy. http://www.psicologia.unipa.it/sciacca2004.htm

Belacchi, C. (2019, May 31-June 1). *Abilità metarappresentive nell'età adulta e negli anziani: confronto tra teoria della mente e competenza definitoria* [Meta-representational skill in adult life and in ageing: comparison between theory of mind and definitional competence]. XII Convegno Nazionale di Psicologia dell'invecchiamento, Chieti, Italy. https://www.sipinvecchiamento.it/Convegni/SIPI_12_ATTI_PESCARA_2019_web.pdf

Belacchi, C., & Artuso, C. (2018). How taxonomic and thematic associations in semantic memory modulate recall in young through old-old adults. *Psychology and Aging, 33*, 1060-1069. https://doi.org/10.1037/pag0000297

Belacchi, C., Artuso, C., & Palladino, P. (2022). Semantic long-term memory and verbal working memory performance: How does their relationship change with age? *Cognitive Development, 62*, 101185. https://doi.org/10.1016/j.cogdev.2022.101185

Belacchi, C., & Benelli, B. (1999). Lo sviluppo della competenza definitoria in soggetti insufficienti mentali e normodotati [Definitional Competence development in low I.Q and typically developing subjects]. *Psicologia Clinica Dello Sviluppo*, *3*, 385-420. https://doi.org/10.1449/574

Belacchi C., Benelli, B. (2007), *Il significato delle parole. La competenza definitoria nello sviluppo tipico e atipico* [The meaning of words. Definitional competence in typical and atypical development]. Il Mulino

Belacchi, C. & Benelli, B. (2017). "A mosaic is not a place where you go to see pictures": On the nature of definitions and the development of definitional skills. *Journal of Applied Psycholinguistics, XVII, 71-91. https://www.torrossa.com/en/resources/an/4202276*

Belacchi, C., Benelli, B. (2021). *Valutare la competenza definitoria. La Scala Co. De. In ambito clinico e nello sviluppo tipico* [Measuring the Definitional Competence. The Co.De Scale in typical and atypical development]. Franco Angeli.

Benelli, B., Belacchi, C., Gini, G., & Lucangeli, D. (2006). 'To define means to say what you know about things': the development of definitional skills as metalinguistic acquisition. *Journal of Child Language*, *33*, 71-97. https://doi.org/10.1017/S0305000905007312

Bianco, F., Lombardi, E., Lecce, S., Marchetti, A., Massaro, D., Valle, A., & Castelli, I. (2021). Supporting children's second-order recursive thinking and advanced ToM abilities: A training study. *Journal of Cognition and Development, 22(4), 1-24.* https://doi.org/10.1080/15248372.2021.1901712

Bianco, F., Lombardi, E., Massaro, D., Castelli, I., Valle, A., Marchetti, A., & Lecce, S. (2019). Enhancing advanced theory of mind skills in primary school: A training study with 7 to 8 year olds. *Infant and Child Development, 28, e2155, 1-16.* https://doi.org/10.1002/icd.2155

Bosacki, S. L., &Astington, J. W. (1999). Theory of mind in preadolescence: Relations between social understanding and social competence. *Social Development, 8, 237-255.* https://doi.org/10.1111/1467-9507.00093

Bowles, R. P., & Salthouse, T. A. (2008). Vocabulary test format and differential relations to age. *Psychology and Aging, 23, 366–376.* https://doi.org/10.1037/ 0882-7974.23.2.366

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition (pp. 373–443). Psychology Press.* https://doi.org/10.1016/B978-0-12-809324-5.01889-7

Cabinio, M., Rossetto, F., Blasi, V., Savazzi, F., Castelli, I., Massaro, D., Valle, A., Nemni, R., Clerici, M., Marchetti, A., & Baglio, F. (2015). Mind-reading ability and structural connectivity changes in aging. *Frontiers, 6, 1808.* https://doi.org/10.3389/fpsyg.2015.01808

Castelli, I., Baglio, F., Blasi, V., Alberoni, M., Falini, A., Sempio, O., Nemni, R., & Marchetti, A. (2010). Effects of aging on mindreading ability through the eyes: An fMRI study. *Neuropsychologia, 48, 2586-2594.* https://doi.org/10.1016/j.neuropsychologia.2010.05.005

Charlton, R. A., Barrick, T. R., Markus, H. S., & Morris, R. G. (2009). Theory of mind associations with other cognitive functions and brain imaging in normal aging. *Psychology and Aging, 24, 338-348.* https://doi.org/10.1037/a0015225

De Rosnay, M., & Hughes, C. (2006). Conversation and theory of mind: Do children talk their way to socio-cognitive understanding? *British Journal of Developmental Psychology, 24, 7-37.* https://doi.org/10.1348/026151005X82901

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender and social experiences in middle childhood. *Child Development, 84, 989-1003.* https://doi.org/10.1111/cdev.12017

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12, 189 –198.* https://doi.org/10.1016/0022-3956(75)90026-6

Gini, G., Benelli, B., & Belacchi, C. (2004). Children's definitional skills and their relations with metalinguistic awareness and school achievement. *School Psychology, 2, 239-267.*

Gutierrez-Clellen, V. F. & De Curtis, L. (1999). Word definition skills in Spanish-speaking children with language impairment. Communication Disorders Quarterly, 21, 23–31. https://doi.org/10.1177/152574019902100104

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging, 28, 826-839.* https://doi.org/10.1037/a0030677

Hughes, C. (2001). Executive dysfunction in autism: Its nature and implications for the everyday problems experienced by individuals with autism. In J. A. Burack, T. Charman, N. Yirmiya & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research (pp. 255–275). Lawrence Erlbaum Associates Publishers.*

Hughes, C. (2011). Social understanding and social lives: From toddlerhood through to the transition to school. Psychology Press.

Hughes, C., Cassidy, B. S., Faskowitz, J., Avena-Koenigsberger, A., Sporns, O., & Krendl, A. C. (2019). Age differences in specific neural connections within the default mode network underlie theory of mind. NeuroImage, 191, 269-277. https://doi.org/10.1016/j.neuroimage.2019.02.024

Karmiloff-Smith A. (1992), *Beyond modularity. A developmental perspective on cognitive development. The MIT Press.*

Lecce, S., Bianco, F., Devine, R. T., & Hughes, C. (2017). Relations between theory of mind and executive function in middle childhood: A short-term longitudinal study. *Journal of Experimental Child Psychology, 163, 69-86.* https://doi.org/10.1016/j.jecp.2017.06.011

Lecce, S., Bottiroli, S., Bianco, F., Rosi, A., & Cavallini, E. (2015). Training older adults on theory of mind: Transfer on metamemory. *Archives of Gerontology and Geriatrics, 60*, 217-226. https://doi.org/10.1016/j.archger.2014.10.001

Lecce, S., Ceccato, I., Bianco, F., Rosi, A., Bottiroli, S., & Cavallini, E. (2017). Theory of Mind and social relationships in older adults: The role of social motivation. *Aging & Mental Health, 21, 253-258.* https://doi.org/10.1080/13607863.2015.1114586

Li, X., Wang, K., Wang, F., Tao, Q., Xie, Y., & Cheng, Q. (2013). Aging of theory of mind: The influence of educational level and cognitive processing. *International Journal of Psychology, 48, 715-727.* https://doi.org/10.1080/00207594.2012.673724

Litowitz, B. (1977). Learning to make definitions. *Journal of Child Language 4, 289–304.* https://doi.org/10.1017/S0305000900001665:

MacPherson, S. E., Phillips, L. H., & Della Sala, S. (2002). Age, executive function, and social decision making: A dorsolateral prefrontal theory of cognitive aging. *Psychology and Aging, 17, 598–609.* https://doi.org/10.1037/0882-7974.17.4.598

Malle B. F (2002), The relation between language and theory of mind in development and evolution. In T. Givòn & B. F. Malle (Eds.), *The evolution of language out of pre-language* (p.265-284*).* Benjamins.

McGhee-Bidlack, B. (1991). The development of noun definitions: A metalinguistic analysis. *Journal of Child Language, 18, 417–34.* https://doi.org/10.1017/S0305000900011132

Miller, S. A. (2012). *Theory of mind: Beyond the preschool years. Psychology Press.*

Nelson K. (1996), Language in cognitive development. The emergence of mediated mind. Cambridge University Press.

O'Hare, A. E., Bremner, L., Nash, M. Happé, F., & Pettigrew, L. M. (2009). A clinical assessment tool for advanced theory of mind performance in 5 to 12 years old. Journal of Autism and Developmental Disorders, 39, 916-928. https://doi.org/10.1007/s10803-009-0699-2

Pardini, M., & Nichelli, P. F. (2009). Age-related decline in mentalizing skills across adult life span. *Experimental Aging Research, 35, 98-106.* https://doi.org/10.1080/03610730802545259

Pezzuti, L., Longobardi, E., Milletti, K., & Ovidi, A. (2011). A study about the theory of mind in primary and secondary aging. *Life Span and Disability, 14, 31-44.*

Phillips, L. H., MacLean, R. D., & Allen, R. (2002). Age and the understanding of emotions: Neuropsychological and sociocognitive perspectives. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 57, 526-530. https://doi.org/10.1093/geronb/57.6.P526

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind. *Behavioral and Brain Sciences, 1, 515–526.* https://doi.org/10.1017/S0140525X00076512.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach. MIT Press.*

Rosch, E., Mervis, C., Gray, W., Johnson, P.M., & Boyes-Bream, P. (1976). Basic objects in natural categories, *Cognitive Psychology, 8*, 382 – 439. https://doi.org/10.1016/0010-0285(76)90013-X

Roselli, I., Valentini, P., Belacchi, C. (2019, May 10-11). *Abilità metarappresentative e metalinguistiche in adolescenti* [Meta-representational and meta-linguistic skills in adolescents] [Poster presentation]. Giornate Clasta X edizione, Verona, Italy. https://www.univr.it/it/iniziative/-/evento/8488?p_auth=QU5slY09

Rosi, A., *Bottiroli, S., Cavallini, E., Bianco, F., & Lecce, S. (2016). Promoting theory of mind in older adults: Does age play a role? Aging & Mental Health*, *20*, 22-28. https://doi.org/10.1080/13607863.2015.1049118

Rossetto, F., Baglio, F., Massaro, D., Alberoni, M., Nemni, R., Marchetti, A., & Castelli, I. (2020). Social cognition in rehabilitation context: Different evolution of affective and cognitive theory of mind in mild cognitive impairment*. Behavioural Neurology, 2020, 1-9.* https://doi.org/10.1155/2020/5204927

Serafin, M., & Surian, L. (2004). Il Test degli Occhi: Uno strumento per valutare la" teoria della mente" [The Eyes Test: An instrument to evaluate "theory of mind"]. *Giornale Italiano di Psicologia, 31, 839-862.* https://doi.org/10.1421/18849

Sinclair, A. (1986). Metalinguistic knowledge and language development. In I. Kurkz, G. W. Shugar & J. H. Danks (Eds.), *Knowledge and language. NorthHolland.*

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-analysis of theory of mind and peer popularity in the preschool and early school years. Child Development, 86, 1159–1174. https://doi.org/10.1111/cdev.12372

Slessor, G., Phillips, L. H., & Bull, R. (2007). Exploring the specificity of age-related differences in theory of mind tasks. *Psychology and Aging, 22, 639-643.* https://doi.org/10.1037/0882-7974.22.3.639

Snow, C. E. (1990). The development of definitional skill. *Journal of Child Language, 17, 697–710.* https://doi.org/10.1017/S0305000900010953

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10, 640-656.* https://doi.org/10.1162/089892998562942

Tomasello, M. (1999). *The cultural origins of human cognition. University Press*

Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In M. Tomasello & E. Bates (Eds.), Language development: The essential readings (p. 111–128). Blackwell Publishing.

Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. Harvard University Press.

Tomasello. M. (2008). Origins of human communication. MIT Press

Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H. (2005) Understanding and sharing intentions: The origins of human social cognition. Behavioral and Brain Sciences, 28, 675–735. https://doi.org/10.1017/S0140525X05000129

Tomasello M, Farrar J. (1986). Joint attention and early language. *Child Development, 57, 1454–1463.* https://doi.org/10.2307/1130423

Valle, A., Massaro, D., Castelli, I., & Marchetti, A. (2015). Theory of mind development in adolescence and early adulthood: the growing complexity of recursive thinking ability. *Europe's Journal of Psychology, 11, 112-124.* https://doi.org/10.5964/ejop.v11i1.829

Wang, J.-j., & Kaufman, A. S. (1993). Changes in fluid and crystallized intelligence across the 20- to 90-year age range on the K-BIT. *Journal of Psychoeducational Assessment, 11(1), 29–37. https://doi.org/10.1177/073428299301100104*

Watson R. (1995). Relevance and definition, Journal of Child Language, 22, 211-222. https://doi.org/10.1017/S0305000900009703

Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology, 15, 728-755.* https://doi.org/10.1080/17405629.2018.1435413

Wellman, H. M., Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75(2), 523-541.* https://doi.org/10.1111/j.1467-8624.2004.00691.x.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13, 103-128.* https://doi.org/10.1016/0010-0277(83)90004-5

World Medical Association. (2014). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *The Journal of the American College of Dentists, 81, 14-18.* https://doi.org/10.1001/jama.2013.281053

Zhang, H., Eppes, A., & Diaz, M. T. (2019). Task difficulty modulates age-related differences in the behavioral and neural bases of language production. *Neuropsychologia, 124, 254–273.* https://doi.org/10.1016/j. neuropsychologia.2018.11.017

https://doi.org/10.17323/jle.2022.13422

Comparing Two Measures of L2 Depth of Vocabulary Knowledge Using their Relationship with Vocabulary Size

Ali Dabbagh [1], **Mostafa Janebi Enayat** [2]

[1] Gonbad Kavous University, Iran

[2] University of Maragheh, Iran

ABSTRACT

Background. This study compared two tests of second language (L2) depth of vocabulary knowledge, namely the word association test (WAT) and vocabulary knowledge scale (VKS), with respect to their associations with vocabulary size. The same relationships were further examined separately for the five word-frequency bands of the vocabulary size test.

Methods. 115 Iranian English as a Foreign Language (EFL) learners who were native speakers of Persian took the WAT, VKS, and Vocabulary Levels Test (VLT). The selected participants were undergraduates who ranged from freshmen to junior and were both male (n=47) and female (n=68) students.

Results. The outcomes of multiple linear regression analyses indicated that: (a) while both measures of vocabulary depth were predictive of the VLT, the WAT had a higher association with the dependent variable; (b) both the WAT and VKS were predictive of the high-frequency vocabulary, with the relationships being more significant for the WAT; (c) the WAT could significantly predict the mid-frequency vocabulary, whereas the VKS had no significant contribution; and (d) while the VKS was significantly associated with the low-frequency vocabulary, the WAT had no significant contribution to the prediction of this level.

Implications. The implications of the findings are interpreted with reference to the suitability of both the WAT and VKS depending on the type of input, expected response, and desired frequency of the target words.

KEYWORDS

vocabulary size, vocabulary depth, Word Association Test (WAT), Vocabulary Knowledge Scale (VKS), Vocabulary Levels Test (VLT)

INTRODUCTION

Vocabulary knowledge has been recognized as one of the most significant components of language learning without which no meaning can be conveyed and understood (Dabbagh & Janebi Enayat, 2019; Duong, 2022; Janebi Enayat & Babaii, 2018; Mathews, 2018; Read, 2004; Roche & Harrington, 2013; Schmitt, 2010; Schmitt, 2014; Schmitt & Schmitt, 2014). Uchihara and Clenton (2022), for instance, found that spoken vocabulary knowledge is significantly correlated with second language (L2) speaking proficiency. Researching vocabulary involves dealing with a multidimensional construct as the nature of this knowledge is perplexing and entails various aspects of form, meaning, and use, each of which encompasses sub-components (Laufer et al., 2004; Nation, 1990; Schmitt, 2014). To grapple with such complexity, a variety of descriptive frameworks have been suggested to systematically categorize the construct of vocabulary knowledge, the most oft-cited of which is the classification of *size* and *depth* (Haastrup & Henriksen 2000; Henriksen, 1999; Qian, 1998; Read 1993; Schmitt, 1999), with the former pertaining to the number of words L2 learners know during a particular stage of the learning process (Nation, 2001) and the latter relating to the quality of word knowledge or how well the L2 learners know a single lexical item (Read, 2000). Depth of vocabulary knowledge, therefore, embodies not only the diction-

Citation: Dabbagh, A., & Janebi Enayat, M. (2022). Comparing Two Measures of L2 Depth of Vocabulary Knowledge Using their Relationship with Vocabulary Size. *Journal of Language and Education, 8*(3), 52-64. https://jle.hse.ru/article/view/13422

Correspondence:
Mostafa Janebi Enayat
m.enayat@maragheh.ac.ir

Received: November 25, 2021
Accepted: June 08, 2022
Published: September 30, 2022

ary definition of a word, but also its semantic network which includes, but is not limited to, paradigmatic and syntagmatic lexical relations (Schoonen & Verhallen, 2008), which refer to the linear relations between two words that could appear in the same sentence (e.g., research-conduct, research-observation, and research-laboratory) and hierarchical relations (e.g., research-science, research-experimental), respectively.

The introduction of these two dimensions of vocabulary knowledge led to the development of some reliable and valid tests to measure them. Vocabulary size, in particular, has attracted more attention in L2 vocabulary research (David, 2008) due to its critical and substantial contribution to effective language use (Alharthi, 2020; Dabbagh & Janebi Enayat, 2019; Derakhshan & Janebi Enayat, 2020; Nguyen & Nation, 2011; Uchihara & Clenton, 2020). Various instruments have been, therefore, developed to examine the L2 learners' vocabulary size. Vocabulary Levels Test (VLT) is perhaps the most widely used measure of vocabulary size (Webb & Sasao, 2013) which was first designed by Nation (1983) and later revised and validated by Schmitt et al. (2001). The test is built upon the five word-frequency levels of 2,000, 3,000, 5,000, and 10,000 that comprise 120 high- and low-frequency target words. Another measure of vocabulary size is the Vocabulary Size Test (VST) designed by Nation and Beglar (2007) and validated by Beglar (2010) that evaluates L2 vocabulary size using fourteen 1,000-word-frequency levels that include 140 lexical items. This test has addressed more word-frequency bands which have made it more comprehensive (Elgort, 2013) and suitable to measure the progress of vocabulary size over time (Beglar, 2010). Another measure of vocabulary size is the Yes/No test format designed by Meara (1992) which, compared to the VLT, was found to be a less effective test (Cameron, 2002). A more recent and modified version of the VLT, known as the New Vocabulary Levels Test (NVLT), has been designed by Webb et al. (2017). The NVLT tests the L2 vocabulary size using the first five 1,000-word-frequency bands.

The tests of vocabulary size, such as the VLT, VST, and NVLT are based on the word-frequency bands. These tests start with high-frequency vocabulary like knowledge of the first 1,000-word-frequency level and end with low-frequency vocabulary, such as the 10,000-word-frequency band in the VLT and the 14,000-word-frequency level in the bilingual and monolingual versions of the VST (Janebi Enayat et al., 2018). Schmitt and Schmitt (2014) reassessed the boundaries and proposed another category. They argued that high-frequency English vocabulary should contain the most frequent 3,000 word families. Additionally, they proposed that the low-frequency vocabulary should be lowered to 9,000-word frequency level and beyond. The authors labelled the vocabulary between high-frequency (3,000) and low frequency (9,000) as the "mid-frequency" vocabulary.

Depth of vocabulary knowledge, however, has gained less attention in language testing as it entails a range of word relations, making it difficult to offer a unified definition for this dimension of word knowledge (Milton, 2009; Schmitt, 2014). In fact, compared to the number of tests developed and validated for measuring vocabulary size, "less progress has been made, both in defining depth as a construct and in developing tests for practical use" (Read, 2007, p. 105). The first is Vocabulary Knowledge Scale (VKS), designed by Pribakht and Wesche (1993) and Wesche and Paribakht (1996), which assesses different stages of vocabulary knowledge ranging from no familiarity with the word to the ability to use it accurately in sentences. However, the instrument which could find its way in almost all of the previous studies on depth of vocabulary was Word Association Test (WAT). Developed and validated first by John Read in 1993, WAT assesses depth of vocabulary knowledge through asking learners to choose only four out of eight responses which may be, in one way or another, related to the cue word. As Read (1993) stated, "it is assumed that learners with a deeper knowledge of the word will be better able to pick the associates (which should represent different aspects of the meaning of the word) than those whose knowledge is more superficial" (p. 395).

Despite the surge of interest in using WAT as a valid measure of depth of vocabulary knowledge, recent studies have revealed that this test might partially score vocabulary size, as well. Previous studies showed significant correlations between scores resulted from WAT and VLT, a measure of vocabulary size (Akbarian, 2010; Huang, 2006; Janebi Enayat et al., 2018; Noro, 2002; Schmitt, 2014). This high interrelationship between the two constructs is in line with the argument proposed by Meara and Wolter (2004) that the two aspects of depth and size are not separate from each other and improvement in vocabulary size results in the development of vocabulary depth as well. The current study has attempted to find the more suitable test of vocabulary depth for university English as a Foreign Language (EFL) learners using this interrelationship between the two aspects of depth and size of vocabulary knowledge, using WAT and VKS as measures of vocabulary depth and VLT and its frequency bands as a measure for vocabulary breadth.

LITERATURE REVIEW

In this section, literature on two main measures of vocabulary depth, i.e., WAT and VKS, and their limitations are briefly reviewed. Then, the previous research on probing the interaction between vocabulary breadth and depth is succinctly reviewed. Highlighting Meara and Wolter's (2004) model as the theoretical background utilized in the present study, the section ends with introducing the research gap and formulating the research questions.

Measures of Vocabulary Depth

The depth dimension of vocabulary knowledge has been measured using a couple of tests, but, compared to the number of measures developed for size aspect, less attempt has been made to both define this dimension as a construct and develop tests to measure it (Read, 2007). Read (2000) classified the different approaches to measure depth of vocabulary knowledge into two main groups: developmental and dimensional. In the former, a scale of measurement is used to describe the stages in vocabulary acquisition. To that end, Paribakht and Wesche (1993) and Wesche and Paribakht (1996) designed VKS. First designed to assess English vocabulary learning in language programs at the University of Ottawa, Canada, this scale measures the different levels of lexical knowledge of particular words being learned in a comprehension-based ESL classroom. This scale is a self-report measure which learners are asked to specify their degree of understanding of individual words on a scale of 1-5. The first three categories of this scale deal with conceptual familiarity with the cue word (from no familiarity to the ability to provide a synonym) and the last two categories involve assessing the productive knowledge of the prompt words by asking to compose a response (category IV: I know this word. It means _____ and category V: I can use this word in a sentence, as follows) (see Figure 1).

Figure 1

Sample of VKS item (Taken from Wesche & Paribakht, 1996)

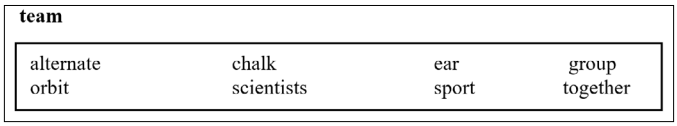
Domestic	
I.	I don't remember having heard this word before.
II.	I have heard this word before but I don't know what it means.
III.	I have heard this word before, and I think it means _____ (synonym or translation)
IV.	I know this word. It means _____ (synonym or translation)
V.	I can use this word in a sentence: _____ (if you do this section, please do section IV)

However, as Qian (1998) argued, VKS assesses only one meaning of the prompt word coupled with its actual use and ignores measuring multiple meanings or associations. Henriksen (1999) further confirmed this argument and noted that VKS only assesses the receptivity or productivity of the target words with no measurement of their different aspects. In addition, Schmitt (2010) listed the following limitations for this scale: 1. the first two stages of the scale are unverified; 2. the underlying knowledge construct are inconsistent, jumping from form-meaning (categories I to IV) to production in context (category V); 3. the intervals between the categories are not consistent; 4. the metalinguistic judgement in categories II (I *think* I know the word) and III (I *know* the word) can be confusing for some learners since they are better at judging what they can do with the words; and, more importantly, 5. the simple sentences examinees write in category V cannot clearly show their productive knowledge of the target word. As Webb (2013) mentioned, in VKS, "it is possible [for test takers] to use a word correctly in a sentence without knowing its meaning" (p.3). In this regard, Zhong (2016) suggested to adapt the test in a way to reach the minimum possible chance for test takers to produce 'neutral' sentences like 'It is beautiful' or 'He is calm'.

The dimensional approach, on the other hand, tries to describe the mastery of various components of different words and considers the mastery of lexical networks of an individual word as important (Read, 1993). To assess such an aspect, WAT was designed and further revised by Read (1993, 1998) which assesses the depth of individual vocabulary knowledge through word association and the relationships between the words in the mental lexicon. This was a developed format of his previous attempt to measure depth of vocabulary through interview procedure in which the learners were asked to pronounce the words, provide an explanation, identify the domain, provide word associations, and suggest other forms of the word (Read, 1998). The first version of the test designed in 1993 includes eight options for each target word of which four were associated with the target word paradigmatically (synonym), syntagmatically (collocation), and analytically (component) (see Figure 2).

Figure 2

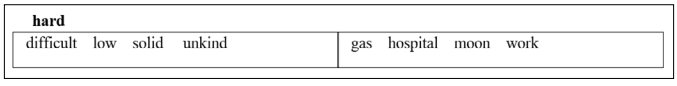
The first version of WAT in 1993



The 1998 version uses two boxes with eight words in each for 40 target words, all of which are adjectives. The examinees should select only four words associated with the target word from the two boxes (see Figure 3). The words in the left box are paradigmatically related to the target word and the ones in the right box are syntagmatically related. To reduce the guessing effect, the patterns of students' responses differ such that three format are possible: two words from the right box and two from the left one; three from the right and one from the left; or three from the left and one from the right.

Figure 3

The 1998 version of WAT



The merit of this test format is in its ability to tap different instances of meaning, collocation, and formulaic language (Schmitt, 2010). Schmitt *et al*. (2011) reported that WAT could be regarded as an appropriate measure of depth of vocabulary since "it is tapping into learners' uncertainty about collocational combinations" (p.118).

Despite its wide use in measuring depth of vocabulary knowledge (e.g., Akbarian, 2010; Atai & Dabbagh, 2010; Dab-

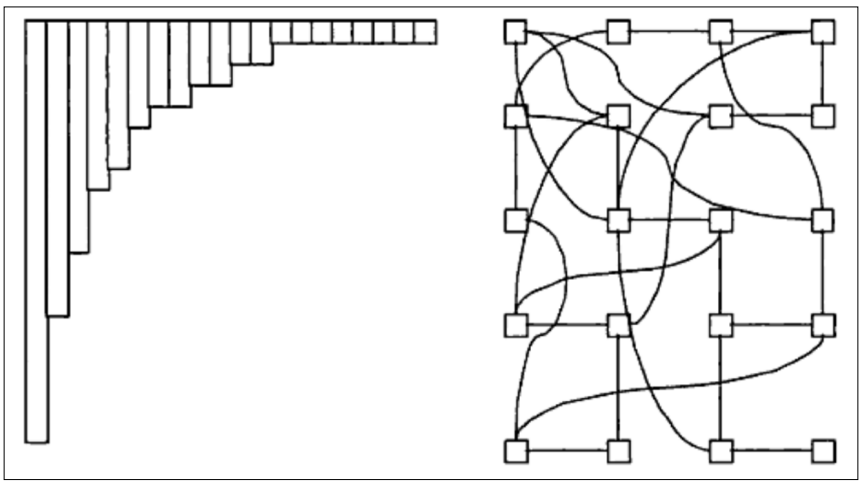
bagh, 2016; Dabbagh & Janebi Enayat, 2019; Janebi Enayat & Babaii, 2018; Janebi Enayat & Derakhshan, 2021; Janebi Enayat et al., 2018; Nassaji 2006, Qian 1999, 2002; Schoonen & Verhallen, 2008, among others), WAT is regarded as a challenging measure of depth of vocabulary knowledge for advanced learners at university level (Greidanus et al., 2005; Greidanus & Neinhuis, 2001; Zhang & Koda, 2017). In addition, different scholars refer to the shortcomings of WAT as a measure of vocabulary depth from various viewpoints. Webb (2013) pointed out that although WAT measures three different aspects of vocabulary depth, namely concept and referents, form and meaning, and collocation, it does not provide separate scores for each of these aspects and it is plausible that two test takers who are actually distinct in their depth of vocabulary dimensions receive the same score without being distinguished in terms of what depth of vocabulary aspect was known by each. Akbarian (2010) highlighted that due to the identification of nouns to be collocated with the adjectives given in the test as target words, the test addresses knowledge of adjectives directly and nouns rather indirectly. Also, adverbs are indirectly focused on in WAT since almost all adverbs are related to their corresponding adjectives (Ishii, 2005). However, measuring depth of knowledge of verbs is taken for granted and not included in the test. In addition, as Milton (2009) and Read (1993, 1998) asserted, WAT is susceptible to guessing due to its receptive multiple-choice format which can threaten the validity of the test. Test takers can easily choose some of the given words on random which can make the score interpretation problematic since scores may not provide a true estimate of the test takers' depth of vocabulary knowledge. As Schmitt *et al.* (2011) asserted in their validation study of WAT, the guessing effect can mostly happen for scores 0-2 and not for scores 3-4 for each item. More specifically, they found that split scores – where test takers achieve 1, 2, or 3 out of the maximum 4 for each item – mostly resulted from no knowledge or partial knowledge of the target word and consequently no clear interpretation can be reached upon for these scores. They also relate guessing in WAT items to its tendency to overestimate the test takers' actual knowledge of the target words and raise the question of whether test takers are successful in guessing even if they have no knowledge of the target words.

The Interconnection between Size and Depth as a Possible Yardstick

Even though being distinct in terms of measurement instrument, the depth and size of vocabulary have been found to be so much inter-related. Nurweni and Read (1999), in a study on the vocabulary knowledge of first-year students in an Indonesian university concluded that the tests of size (word translation test) and depth of vocabulary (WAT) correlated highly with each other (r = .62). Qian (1999) explored this issue and found a significant correlation of .78 between the VLT and WAT scores among 44 Korean and 33 Chinese speakers. Henriksen (1999) further argued that "an understanding of the relations among the items is a prerequisite for a more precise understanding of each individual item" (p. 313). This interconnection between the two dimensions of size and depth of vocabulary was supported by Meara and Wolter (2004) who believed that "vocabulary size is not a feature of individual words: rather it is a characteristic of the test taker's entire vocabulary" (p. 87). These two scholars proposed two different models for the interconnection between size and depth of vocabulary (see Figure 4). While in the first one (the left hand diagram) vocabulary size and depth are not intrinsically interrelated and adding more lexical items does not develop the whole lexicon, in the second model (the right hand diagram), an increase in vocabulary size could develop the lexical network (vocabulary depth) as well. This relationship was also approved by Ishii and Schmitt (2009) who contended the two aspects are interconnected such that one dimension would be incomplete without the other.

Figure 4

Two ways of looking at the relation between vocabulary size and depth. From "V-Links: beyond vocabulary depth," by P. Meara and B. Wolter, in D. Albrechtsen, K. Haastrup, and B. Henriksen (Eds.), Writing and vocabulary in foreign language acquisition (p. 89), 2004, Museum Tusculanum. Press



Many studies have reported the interconnection between these two dimensions of vocabulary size and depth, as measured by the VLT and WAT, respectively (e.g., Akbarian, 2010; Gyllstad, 2007; Huang, 2006; Janebi Enayat et al., 2018; Qian, 2002; D. Zhang, 2012). This interconnection, however, does not mean that the test takers' scores on the VLT could show both size and depth of vocabulary knowledge because tests of receptive vocabulary size intend to measure form/meaning recognition knowledge, and not vocabulary depth which is assessed using word associations tasks. Put it more simply, the relationship between these two dimensions could possibly mean that they are related to the same construct and, therefore, should not be seen as separate aspects (Vermeer, 2001). Another interpretation is that tests of depth of vocabulary knowledge, such as WAT, are not actually tests of vocabulary depth; they are rather size tests "masquerading as depth tests" (Akbarian, 2010, p. 400). This claim was also backed by Milton (2009) who asserted that the associative format to measure depth of vocabulary is not successful in measuring this vocabulary construct for the main reason that this format is incapable of tapping into the quality of association the test takers make.

The correlation between size and depth of vocabulary knowledge has been reported to be unclear for lower and higher frequency words. While there seems to be little difference between these two dimensions for higher frequency words, a gap has been reported between these aspects of vocabulary for lower frequency words (Schmitt, 2014). Shimamoto (2000), Noro (2002), and Henriksen (2008) for instance, found the relationship to be weaker for learners who had larger vocabularies and higher language proficiency.

In the present study, the proposed model of Meara and Wolter (2004) and the interconnection between the two dimensions of vocabulary size and depth were utilized as a yardstick to identify the most suitable test of vocabulary depth. Additionally, the nature of this relationship was probed for higher and lower word-frequency levels of the vocabulary size test. Therefore, the following research questions were formulated in this study: (1) Which measure of vocabulary depth has the highest predictive ability for L2 vocabulary size? (2) How is the predictive ability of the two measures of vocabulary depth in L2 vocabulary size different for high and low word-frequency bands?

METHODS

In this section, the demographic information of the participants as well as the instruments used for data collection are explained. The steps followed for data collection and analysis are also described.

Participants

A sample of 115 intermediate EFL undergraduate learners, who were all native speakers of Persian, was selected based on the results of the quick Oxford Placement Test (2004) out of 234 Iranian undergraduate students of English Language Teaching and English Language and Literature. Accordingly, the participants who scored between 30 and 47 in the test, i.e., B1 and B2 according to the Common European Framework of Reference (CEFR), were selected. The selected participants ranged from freshmen to junior who were both male (n=47) and female (n=68) students with the age range of 18 to 25. The reason for selecting this sample is that based on the nature of the study, participants should have a good mental lexicon in terms of quality and quantity of word knowledge and an acceptable command of English.

Instruments

Oxford Quick Placement Test (OQPT, 2004)

To homogenize the participants in terms of the proficiency level, this test was administered and the ones with intermediate level of English language proficiency were selected. The test, which was developed by Oxford University Press and University of Cambridge Local Examinations Syndicate, consists of 60 multiple-choice items to which participants were to answer in 30 minutes. According to the OQPT scoring system, the participants who scored between 26 and 45 were determined as intermediate language level. A high validity and a reliability close to .90 was reported by Geranpayeh (2003) for this version of the test.

WAT-Test of Dimensional Aspect of Vocabulary Depth

Developed by Read (1993), WAT measures the depth of vocabulary knowledge of the participants. The test is a list of 40 prompt words each of which consists of one stimulus word, which is an adjective, followed by a list of eight words in two boxes of four words. The left and right boxes consist of the synonymous words and collocations of the stimulus words, respectively. The participants should choose four words that are related to the prompt word semantically. The four related words have been selected to represent three semantic relations, namely paradigmatic, syntagmatic and analytic (Read, 1993). Read (1995) reported its reliability (KR-20, N=94) as .93 and Nassaji (2006) and Qian (2002) found its split half reliability to be .89.

VKS- Test of Developmental Aspect of Vocabulary Depth

Developed originally by Paribakht and Wesche (1993), VKS was used to find out the participants' self-perceived level of developmental aspect of depth of vocabulary knowledge. Participants should indicate their level of knowledge about the target words on a Likert scale ranging from total unfa-

miliarity to the ability to use the words in context. The instrument enjoys a high reliability estimate of .89 for content words and .82 for discourse connectives as reported by Paribakht and Wesche (1997).

As VKS is a tool which in theory can be used with any set of words and since the aim of the current study is to compare VKS and WAT, the same prompt words in the latter test were utilized as the cue words for the former.

VLT-Test of Vocabulary Size

Designed by Nation (1983) as a measure of breadth of vocabulary, this test "provides a profile of a learner's vocabulary" (p.58) in terms of levels of frequency (2000-, 3000-, 5000-, and 10,000-word-frequency levels) with large samples of words from different frequency levels. In other words, "[the test scores] obtained from VLT were treated as the variable of size of vocabulary knowledge" (Akbarian, 2010, emphasis added). The test has been validated and revised by many scholars since its first format (e.g., Ishii & Schmitt, 2009; Schmitt et al., 2001; Xing & Fulcher, 2007). Version 2 of this test, which was revised and validated by Schmitt *et al.* (2001), is employed in this study. In this version, the participants were given 10 groups of words in each frequency level. Each group consists of 6 cue words that should be matched with 3 definitions (see Figure 5). The test has been reported as reliable with a Cronbach alpha of .96 (Akbarian, 2008) and .81 (Schmitt et al., 2001).

Figure 5.

A VLT sample item (Taken from Schmitt et al., 2001)

1.	accident	
2.	debt	…. loud deep sound
3.	fortune	…. something you must pay
4.	pride	…. having a high opinion of yourself
5.	roar	
6.	thread	

Although Nation and Beglar's (2007) Vocabulary Size Test (VST) has been claimed to be a more comprehensive measure of breadth of vocabulary than the VLT, the present study utilized the latter for the reason that the four-option multiple-choice format of VST is subject to guessing effect (Gyllstad *et al.,* 2015), which may lead to the overestimation of test scores over and above the six-option matching format in VLT (Stewart, 2014; Stewart & White, 2011). Moreover, VLT is the widely used measure of breadth of vocabulary among researchers (e.g., Abdullah et al., 2013; Akbarian, 2010; Alavi & Akbarian, 2012; Baba, 2009; Dabbagh, 2016; Dabbagh & Janebi Enayat, 2019; Janebi Enayat & Derakhshan, 2021; Janebi Enayat et al., 2018; Qian, 2002; Webb & Sasao, 2013; Zhang & Anual, 2008).

Procedure and Data Analysis

First, the Oxford Quick Placement Test was administered to determine participants' proficiency level and select the intermediate ones. Then, to measure the participants' depth and size of vocabulary knowledge, WAT, VKS and VLT were administered with a one-week time interval for each test to prevent sensitization of students to the purpose of the research and control the testing effect. While administering the WAT, the participants were encouraged to give as many answers as they could, even if they would not be sure whether the given answers were correct or not (Read, 1993). As for the VLT, the participants were required not to follow the guessing strategy for the words they did not know, but they were suggested to find the answer if they thought they might know it. The time allotted for each test was 30 to 45 minutes. The WAT, VKS, and VLT papers of the participants were scored following the criteria established by Nassaji (2006), Wesche and Paribakht (1996), and Schmitt *et al.* (2001), respectively. Multiple linear regression analyses were run using SPSS version 23.0 to find the contribution of WAT and VKS to VLT and the extent that the high and low word-frequency bands were predicted by the two tests of vocabulary depth.

RESULTS

Descriptive and Reliability Statistics

Table 1 represents a general profile of the descriptive statistics of the participants' scores on the WAT, VKS, VLT, and the four word-frequency bands of the VLT. As the data analyzed below shows (see Table 1), the participants' scores on the three administered vocabulary tests and the sub-tests of the VLT enjoyed appropriate Cronbach's alpha reliability estimate, which means these tests enjoy high reliability for the sample of the present study.

Before running multiple regression analyses, the correlations among the variables were calculated. The results of Shapiro-Wilk test indicated that except for the scores on the VKS and VLT, the scores on the other WAT and the four sub-tests of the VLT were not normally distributed ($p > .05$). Spearman correlation coefficients were calculated for the sets of scores, the results of which are provided in Table 2. It shows that the correlations among all the variables were significant ($p < .05$) and the correlations between the VKS and WAT, as the predictor variables were also significant ($p < .05$). However, multicollinearity, i.e., correlation of independent variables in a regression model (Field, 2009), was not a concern as the tolerance values were less than 0.40 and the variance inflation factors (VIFs) were less than 2.5 (Field, 2009).

Table 1

Descriptive statistics for the participants' scores on each test and sub-test.

Test	MPS	Min.	Max.	Mean	SD	α
WAT	100	18	65	46.99	9.84	.83
VKS	200	75	182	138.22	20.30	.86
VLT	120	33	85	64.14	11.15	.89
VLT 2,000	30	20	30	27.49	2.38	.74
VLT 3,000	30	10	29	21.94	4.43	.76
VLT 5,000	30	1	24	12.75	5.23	.79
VLT 10,000	30	0	6	1.94	1.49	.81

Note: N = 115. MPS = Maximum possible score; SD = standard deviation; α = Cronbach's alpha.

Table 2

Spearman correlation coefficients among the vocabulary depth and size tests and sub-tests

Test	WAT	VKS	VLT	VLT 2K	VLT 3K	VLT 5K	VLT 10K
WAT	-						
VKS	.313**	-					
VLT	.433**	.430**	-				
VLT 2K	.396**	.363**	.671**	-			
VLT 3K	.454**	.397**	.832**	.522**	-		
VLT 5K	.337**	.283**	.895**	.467**	.586**	-	
VLT 10K	.231*	.422**	.666**	.428**	.375**	.640**	-

*Note: N = 115. K = 1,000. **p < .01 **p < .05*

Predictive Ability of WAT and VKS in VLT

To answer the first research question, the contribution of the participants' WAT and VKS scores to VLT scores was examined through multiple linear regression analysis (using the stepwise method). The results, as shown in Table 3, revealed that two models emerged for this association. The first model in which only the WAT was entered as the predictor variable could explain about 23% of the variance in the VLT (*F (1,113) = 33.565, p < .001,* R^2 = .229). The second model where both WAT and VKS were entered as the explanatory variables could explain 29% of the VLT performance (F (2,112) = 23.052, p < .001, R^2= .292). In other words, the addition of the VKS scores could provide an additional 6% of the predictive power which was a significant change (p < .01). This shows that WAT had a more significant association with the test of vocabulary size compared to the VKS test. Put it simply, the receptive format of vocabulary depth was more predictive of the scores on the test of vocabulary size than the productive format.

The standardized beta weights also reaffirmed the strength of the association between the scores on the WAT and VLT in the first (β = .479, t = 5.794, p < .001) and second (β = .355, t = 4.005, p < .001) models. The VKS, however, made a less contribution to the prediction of the VLT scores (β = .279, t = 3.146, p < .01).

Predictive Ability of WAT and VKS in High and Low Frequency Vocabulary of VLT

The second research question of the current study investigated the extent that the WAT and VKS scores could predict the high and low word-frequency bands of the VLT. A series of multiple linear regressions (using the stepwise method) were run for this purpose. The results (see Table 4) indicated that, for the 2,000-word-frequency band of the VLT, two models emerged. In the first model, only the WAT was entered as the predictor variable which could explain 15.5% of the variance in this sub-test of the VLT (*F (1,113) = 20.781, p < .001,* R^2 = .155). The second model in which both WAT and VKS were entered as the predictor variables could explain 19.5% of this word-frequency band of the VLT (F (2,112) = 13.580, p < .001, R^2= .195). The addition of the VKS scores could, therefore, add 4% to the predictive power which was a significant change (p < .05). Similar results were found for the 3,000-word-frequency band as two models emerged for this dependent variable in the first of which only the WAT

was entered capable of explaining 24% of the variance in the scores of this sub-test (F (1,113) = 35.722, p < .001, R^2= .240). In the second model where both WAT and VKS were entered as the explanatory variables, the predictive power was 31.5% (F (2,112) = 25.752, p < .001, R^2= .315), indicating that the additional variance explained by the insertion of the VKS was about 7% which was statistically significant (p < .01). As for the 5,000-word-frequency band, one model emerged in which WAT was the only predictor variable capable of explaining 14% of the variance in the scores obtained on this sub-test of the VLT (F (1,113) =8.575, p < .001, R^2= .141). In contrast, in the one model appeared for the 10,000-word-frequency band of the VLT, it was the VKS scores which could significantly provide a similar prediction for the dependent variable (F (1,113) = 19.075, p < .001, R^2= .144). The results, therefore, indicated that the WAT was more associated with the high- and mid-frequency vocabulary size which are measured through the 2,000-, 3,000-

Table 3

Multiple regression analyses for vocabulary depth measures in vocabulary size.

	R	*R²*	*ΔR²*	Unstandardized		Standardized
				B	SE *B*	β
Model 1	.479	.229***				
Constant				38.664	4.493	
WAT				.542	.094	.479***
Model 2	.540	.292***	.063**			
Constant				24.044	6.349	
WAT				.402	.100	.355***
VKS				.153	.049	.279**

***p < 0.01, ***p < 0.001.*

Table 4

Multiple regression analyses for vocabulary depth measures in word-frequency levels of the VLT

Dependent	Predictor	*R*	*R²*	*ΔR²*	Unstandardized		Standardized
					B	SE *B*	β
VLT 2K	Model 1	.394	.155***				
	Constant				23.009	1.005	
	WAT				.095	.021	.394***
	Model 2	.442	.195***	.040*			
	Constant				20.516	1.447	
	WAT				.072	.023	.296**
	VKS				.026	.011	.223*
VLT 3K	Model 1	.490	.240***				
	Constant				11.568	1.774	
	WAT				.221	.037	.490***
	Model 2	.561	.315***	.075**			
	Constant				5.211	2.483	
	WAT				.160	.039	.355***
	VKS				.067	.019	.305**
VLT 5K	Model 1	.376	.141***				
	Constant				3.372	2.224	
	WAT				.200	.046	.376***
VLT 10K	Model 1	.380	.144***				
	Constant				-1.911	.893	
	VKS				.028	.006	.380***

**p < 0.05, **p < 0.01, ***p < 0.001.*

(high-frequency words), and 5,000-word-frequency vocabulary (mid-frequency vocabulary), while the VKS was more linked with the 10,000-word-frequency band that relates to the low-frequency vocabulary size.

Appraisal of the standardized beta further confirmed the significant associations between the WAT scores and the 2,000-word-frequency band ($\beta = .394$, $t = 4.559$, $p < .001$), the 3,000-word-frequency level ($\beta = .490$, $t = 5.977$, $p < .001$), and the 5,000-word-frequency level ($\beta = .376$, $t = 4.310$, $p < .001$). The links between the VKS scores and the 2,000-word-frequency band ($\beta = .223$, $t = 2.354$, $p < .05$) as well as the 3,000-word-frequency level ($\beta = .305$, $t = 3.497$, $p < .01$) were comparatively less significant. In contrast, while the WAT performance was the only variable associated with the 5,000-word-frequency level ($\beta = .376$, $t = 4.310$, $p < .001$), the VKS was the only format which could be linked with the 10,000-word-frequency band ($\beta = .380$, $t = 4.368$, $p < .001$).

DISCUSSION

The current study was an attempt to identify the most suitable measure of vocabulary against the yardstick of associations with VLT, a measure of vocabulary size. The findings of multiple linear regression analyses for the scores of 115 EFL students indicated that the WAT was more significantly associated with the VLT scores, particularly the high- and mid-frequency bands. The VKS, however, had a comparatively weaker contribution to the prediction of the VLT scores, but its prediction of the low-frequency band of this test was unique.

The findings indicated that the interconnection between size and depth, as two aspects of vocabulary knowledge, was strong, as measured through the WAT and VLT, supporting previous studies (Akbarian, 2010; Gyllstad, 2007; Henriksen, 2008; Milton, 2009; Zareva, 2005). For instance, Akbarian (2010) used regression analysis and reported that WAT could predict the variance in the VLT. This study also found that the links between the higher frequency words of the VLT and WAT were stronger than the 10,000-word-frequency band. This could somehow support Schmitt's (2014) conclusion that for higher levels of vocabulary size "there is often little difference between size and a variety of depth measures" while this association is weak for lower frequency bands of the VLT where "there is often a gap between size and depth, as depth measures lag behind the measures of size" (p. 941). Noro (2002) and Henriksen (2008) further reported a less significant correlation between the VLT and WAT for lower frequency words. The strong association between the two tests could be justified with reference to the findings of Meara and Wolter (2004) who reported that an increase in vocabulary size could lead to an increase in vocabulary depth, particularly for lower levels of language proficiency.

In a more recent study, Janebi Enayat and Amirian (2020) found that the VLT and WAT are significantly correlated, particularly for lower-intermediate students. The association, however, was not high for advanced L2 learners. Dabbagh and Janebi Enayat (2019) further found high correlations between the size and depth aspects of vocabulary knowledge, as measured through the VLT and WAT, respectively. This means that the two dimensions could jointly contribute to the overall L2 language proficiency. For instance, Janebi Enayat and Derakhshan (2021) investigated the contribution of vocabulary size and depth to L2 speaking ability using the VLT and WAT and found that the two aspects could jointly predict the L2 oral proficiency.

The results further indicated that the prediction of the VLT was mainly made by the WAT while the VKS contributed to the prediction of the VLT scores less significantly. This could be due to the different task format of the WAT, which employs matching items, while the VKS uses a scale that indicates knowledge subjectively. The objective matching format of the WAT is more compatible with the matching type of the VLT which both reduce the guessing effect (Stewart, 2014). Therefore, the students' score on the WAT could be a more precise indication of their depth of vocabulary knowledge than the VKS which is more subjective. The findings also showed a lower power of VKS than that of WAT in predicting VLT. This finding implies that WAT can be regarded as a measure of depth of vocabulary that is more influenced by the size dimension of vocabulary knowledge, and hence, according to Meara and Wolter's (2004) model, it might be regarded as a better measure of depth of vocabulary in comparison with the VKS.

The results for the second research question revealed that, surprisingly, while the WAT was more predictive for the high- and mid-frequency vocabulary (Schmitt & Schmitt, 2014), for the 10,000-word-frequency band of the VLT, the VKS was the only predictor variable. This can be further discussed in that the partial receptive/productive nature of the VKS can better picture knowledge of less frequent vocabulary compared to the WAT which is only receptive. As it was mentioned previously rather implicitly, the first three columns of VKS measure receptive aspects of depth of vocabulary knowledge and the other two columns focus on the productive aspect. This special feature of VKS makes it more suitable to measure depth of vocabulary both receptively and productively. On the contrary, WAT is mainly a receptive measure of vocabulary depth dealing with making associations among the given words. The difference between receptivity and productivity of these two depth of vocabulary tests can be regarded as the cause of their distinction in regression analysis results. However, there is no doubt that low-frequency vocabulary would be recognized to a larger extent than being mastered productively because these words are supposed to be the most difficult ones in the test. This could be justified further by taking into account the fact that productivity is not always about making complex meaningful

sentences and even writing a simple synonym would make a test productive. Additionally, the fourth scale (see Figure 1), which is identified as the productive part of the scale, asks the students to provide a synonym or simply a translation for the target word. Consequently, the students could just write a translation for the target word which could be easier for the students than encountering the target word in a receptive test which requires knowing some other words in the list of options. What is more, due to the format of the WAT where the students must choose four words from a list of eight words for a target word (see the instruments sections for further information on the patterns of responses), the four responses for the target word are dependent on each other, which could make this "receptive" format more difficult than the fourth scale of the VKS where only a translation would suffice to inform the examiner that the student may have a partial knowledge of the target word. As a result, the probability of knowing a low-frequency word could be more on the VKS than the WAT. This provides empirical support for Read's (2004) proposal calling for distinguishing among different aspects of depth of vocabulary with different measures.

Taking the overall results into account, it can be claimed that although WAT was shown to be more predictive as a measure of vocabulary size, and hence a better measure for depth of vocabulary than VKS in this regard. Each of these tests should be used depending on the purpose of measurement, i.e., whether to measure receptive or productive aspects of depth of vocabulary. Moreover, for tapping less frequent aspects of vocabulary depth, the VKS would be a more suitable option as it has the expected response of only providing a synonym or a translation for the target word.

CONCLUSION

The current study compared WAT and VKS in order to find the most appropriate measure of vocabulary depth via comparing their power to predict VLT scores, as a measure of vocabulary size. It can be concluded that although the WAT scores explain the variance in the VLT scores to a larger extent and could be, therefore, considered a more suitable test of vocabulary depth when we consider the association of size and depth as a yardstick, the VKS should also be seen as a more subjective test of vocabulary depth that could tap into the more productive aspect of this dimension of vocabulary knowledge. The results shed light on the difference between WAT and VKS reporting a low correlation between the two which signifies that they cannot be used for research and instruction purposes interchangeably. Rather they should be used for the purposes which correspond to the nature of their test item structure. In other words, vocabulary researchers can use VKS when they are exploring the role of depth of vocabulary in speaking and writing performance, as productive skills, especially if the focus of the investigation is on less frequent words. Also, WAT can be used in probing the association between reading and/or listening comprehension, as receptive skills, and vocabulary depth. With this specification of the use of measures of depth of vocabulary knowledge, more precise results might be achieved in future vocabulary studies.

The results and conclusion of the present study need to be interpreted with caution as there were some limitations which lead to some suggestions for further research. First of all, similar to previous quantitative studies on WAT, VKS, and VLT, this investigation was based on correlational design and quantitative data. Further qualitative can deepen our understanding of the learners' perceptions and processes in answering items of these tests. Second, this study used the VLT, which is not a comprehensive test of vocabulary size. Future studies can be conducted using different measures of vocabulary size, such as Peabody Picture Vocabulary Test (PPVT) and Vocabulary Size Test (VST) to find about the overlap of depth of vocabulary knowledge with other aspects of vocabulary size. Third, this study focused on one language proficiency level to control the effect of this variable and homogenize the students. The interaction among WAT, VKS, and VLT can be assessed benefiting EFL learners from different proficiency levels.

DECLARATION OF COMPETING INTEREST

None declared.

REFERENCES

Abdullah, K. I., Puteh, F., Azizan, A. R., Hamdan, N. N., & Saude, S. (2013). Validation of a controlled productive Vocabulary Levels Test below the 2000-word level. *System, 41(2), 352–364.* https://doi.org/10.1016/j.system.2013.03.005

Akbarian, I. (2008). *The role of vocabulary knowledge in predicting performance on reading comprehension item types [Unpublished doctoral dissertation]. University of Tehran, Iran.*

Akbarian, I. (2010). The relationship between vocabulary size and depth for ESP/EAP learners. *System, 38*(3), 391–401. https://doi.org/10.1016/j.system.2010.06.013

Alavi, S. M., & Akbarian. I. (2012). The role of vocabulary size in predicting performance on TOEFL reading item types. *System, 40*(3), 376–385. https://doi.org/10.1016/j.system.2012.07.002

Alharthi, T. (2020). Investigating the relationship between vocabulary knowledge and FL speaking performance. *International Journal of English Linguistics, 10(1), 37–46.* https://doi.org/10.5539/ijel.v10n1p37

Atai, M. R. & Dabbagh, A. (2010). Investigating vocabulary depth and semantic set in EFL learners' vocabulary use in writing. *TELL Journal, 4(2), 27–49.*

Baba, K. (2009). Aspects of lexical proficiency in writing summaries in a foreign language. *Journal of Second Language Writing, 18*(3), 191–208. https://doi.org/10.1016/j.jslw.2009.05.003

Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing, 27(1), 101–118.* https://doi.org/10.1177%2F0265532209340194

Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research, 6(2), 145–173.* https://doi.org/10.1191%2F1362168802lr103oa

Dabbagh, A. (2016). The predictive role of vocabulary knowledge in listening comprehension: depth or breadth? *International Journal of English Language and Translation Studies, 4(3), 1–13.*

Dabbagh, A., & Janebi Enayat, M. (2019). The role of vocabulary breadth and depth in predicting second language descriptive writing performance. *Language Learning Journal, 47(5), 575–590.* https://doi.org/10.1080/09571736.2017.1335765

David, A. (2008). Vocabulary breadth in French L2 learners. *The Language Learning Journal, 36(2), 167–180.* https://doi.org/10.1080/09571730802389991

Derakhshan, A., & Janebi Enayat, M. (2020). High- and mid-frequency vocabulary size as predictors of Iranian university EFL students' speaking performance. *Iranian Journal of English for Academic Purposes, 9(3), 1–13.*

Duong, T. M. (2022). Insights into ESP Vocabulary Learning Strategies Used by Vietnamese Tertiary Students. *Journal of Language and Education, 8*(1), 38–49. https://doi.org/10.17323/jle.2022.10924

Elgort, I. (2013). Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing, 30(2), 253–272.* https://doi.org/10.1177%2F0265532212459028

Field, A. (2009). *Discovering statistics using SPSS (3*rd ed.). Sage Publications.

Geranpayeh, A. (2003). A quick review of the English Quick Placement Test. *UCLES Research Notes, 12, 8–10.*

Greidanus, T., & Nienhuis, L. (2001). Testing the quality of word knowledge in a second language by means of word associations: Types of distracters and types of associations. *Modern Language Journal, 85*(4), 567–577. https://doi.org/10.1111/0026-7902.00126

Greidanus, T., Beck, B., & Wakely, R. (2005). Testing the development of French word knowledge by advanced Dutch and English-speaking learners and native speakers. *Modern Language Journal, 89*(2)*, 221–233.* https://doi.org/10.1111/j.1540-4781.2005.00276.x

Gyllstad, H. (2007). *Testing English collocations. Media-Tryck, Lund University.*

Gyllstad, H., Vilkaite, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple choice formats: Issues with guessing and sampling rates. *ITL International Journal for Applied Linguistics, 166*(2), 278–306. https://doi.org/10.1075/itl.166.2.04gy1

Haastrup, K., & Henriksen. B. (2000). Vocabulary acquisition: acquiring depth of knowledge through network building. *International Journal of Applied Linguistics, 10*(2), 221–240. https://doi.org/10.1111/j.1473-4192.2000.tb00149.x

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*(2), 303–317. https://doi.org/10.1017/S0272263199002089

Henriksen, B. (2008). Declarative lexical knowledge. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Vocabulary and writing in a first and second language* (pp. 22–66). *Palgrave Macmillan.*

Huang, H.-F. (2006). *Breadth and depth of English vocabulary knowledge: Which really matters in the academic reading performance of Chinese university students?* [Unpublished master's thesis]. McGill University.

Ishii, T. (2005). *Diagnostic tests of vocabulary knowledge for Japanese learners of English* [Unpublished doctoral dissertation]. University of Nottingham.

Ishii, T., & N. Schmitt. (2009). *Developing an integrated diagnostic test of vocabulary size and depth. RELC Journal, 40*(5), 5–22. https://doi.org/10.1177%2F0033688208101452

Janebi Enayat, M., & Amirian, S. M. R. (2020). The relationship between vocabulary size and depth for Iranian EFL learners at different language proficiency levels. *Iranian Journal of Language Teaching Research, 8(2), 97–114.* https://dx.doi.org/10.30466/ijltr.2020.120891

Janebi Enayat, M., & Babaii, E. (2018). Reliable predictors of reduced redundancy test performance: The interaction between lexical bonds and test takers' depth and breadth of vocabulary knowledge. *Language Testing, 35(1), 121–144.* https://doi.org/10.1177/0265532216683223

Janebi Enayat, M., & Derakhshan, A. (2021). Vocabulary size and depth as predictors of second language speaking ability. *System, 99, 1–15.* https://doi.org/10.1016/j.system.2021.102521

Janebi Enayat, M., Amirian, S. M. R., Zareian, G., & Ghaniabadi, S. (2018). Reliable measure of written receptive vocabulary size: Using the L2 depth of vocabulary knowledge as a yardstick. *Sage Open, 8(1), 1–15.* https://doi.org/10.1177/2158244017752221

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing*, *21*(2), 202–226. https://doi.org/10.1191%2F0265532204lt277oa

Mathews, J. (2018). Vocabulary for listening: Emerging evidence for high and mid-frequency vocabulary knowledge. *System*, *72*, 23–36. https://doi.org/10.1016/j.system.2017.10.005

Meara, P., & Wolter, B. (2004). V-Links: beyond vocabulary depth. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Writing and vocabulary in foreign language acquisition* (pp. 85–96). Museum Tusculanum Press.

Milton, J. (2009*). Measuring second language vocabulary acquisition. Multilingual Matters.*

Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *The Modern Language Journal, 90*(3), *387–40.* https://doi.org/10.1111/j.1540-4781.2006.00431.x

Nation, I.S.P. (1983). Testing and teaching vocabulary. *Guidelines, 5*, 12–25.

Nation, I.S.P. (1990). *Teaching and learning vocabulary.* Newbury House.

Nation, I.S.P. (2001). *Learning vocabulary in another language.* Cambridge University Press.

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Japan Association for Language Teaching, 31*(7), *9–12.*

Nguyen, L.T.C., & Nation, I.S.P. (2011). A bilingual vocabulary size test of English for Vietnamese learners. RELC Journal, 42(1), 86–99. https://doi.org/10.1177%2F0033688210390264

Noro, T. (2002). The roles of depth and breadth of vocabulary knowledge in reading comprehension in EFL. *ARELE: Annual Review of English Language Education in Japan*, *13*, 71–80. https://doi.org/10.20581/arele.16.0_141

Nurweni, A. & Read, J. (1999). The English vocabulary knowledge of Indonesian university students. *English for Specific Purposes, 18*(2), 161–175. https://doi.org/10.1016/S0889-4906(98)00005-2

Paribakht, T. S., & Wesche, M. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal, 11*(1), 9–29. https://doi.org/10.18806/tesl.v11i1.623

Paribakht, T.S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin *(Eds.), Second language vocabulary acquisition: A rationale for pedagogy (pp.174–200). Cambridge University Press.*

Qian, D. (1998). *Depth of vocabulary knowledge: assessing its role in adults' reading comprehension in English as a second language.* [Unpublished doctoral dissertation]. University of Toronto, Canada.

Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review, 56*(2), 283–307. https://doi.org/10.3138/cmlr.56.2.282

Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective*. Language Learning, 52*(3), 513–536*.* https://doi.org/10.1111/1467-9922.00193

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing, 10(3), 355–371.* https://doi.org/10.1177%2F026553229301000308

Read, J. (1995). Validating the word association format as a measure of depth of vocabulary knowledge. *Paper presented at the 17th Language Testing Colloquium, Long Beach, CA.*

Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A.J. Kunnan (Ed.), *Validation in language assessment (pp. 41–60)*. Lawrence Erlbaum.

Read, J. (2000). *Assessing vocabulary.* Cambridge University Press.

Read, J. (2004). Plumbing the depth: how should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language, (pp. 209–227)*. Benjamins.

Read, J. (2007). Second language vocabulary assessment: Current practices and new directions. *International Journal of English Studies, 7(*2), 105–125. https://doi.org/10.6018/ijes.7.2.49021

Roche, T., & Harrington, M. (2013). Recognition vocabulary knowledge as a predictor of academic performance in an English as a foreign language setting. *Language Testing in Asia*, *3*(12), 1–13. https://doi.org/10.1186/2229-0443-3-12

Schmitt, N. (1999). The relationship between TOEFL vocabulary items and meaning, association, collocation and word-class knowledge. *Language Testing, 16(2), 189–216.* https://doi.org/10.1177%2F026553229901600204

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual. Palgrave Macmillan.*

Schmitt, N. (2014). Size and depth of vocabulary knowledge: what the research shows. *Language Learning, 64*(4), 913–951. https://doi.org/10.1111/lang.12077

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching, 47(4), 484–503.* https://doi.org/10.1017/S0261444812000018

Schmitt, N., Ching Ng, J. W., & Garras, J. (2011). The word associates format: Validation evidence. *Language Testing, 28(1), 105–126.* https://doi.org/10.1177%2F0265532210373605

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the vocabulary levels test. *Language Testing, 18(1), 55–88.* https://doi.org/10.1177%2F026553220101800103

Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing, 25(2), 211–236.* https://doi.org/10.1177%2F0265532207086782

Shimamoto, T. (2000). An analysis of receptive vocabulary knowledge: Depth versus breadth. *The Japan-British Association for English Teaching, 4, 69–80.*

*S*tewart, J. (2014). Do multiple-choice options inflate estimates of vocabulary size on the VST? *Language Assessment Quarterly, 11*(3), 271–282. https://doi.org/10.1080/15434303.2014.922977

Stewart, J., & White, D. A. (2011). Estimating guessing effects on the Vocabulary Levels Test for differing degrees of word knowledge. *TESOL Quarterly, 45(2), 370–380.* https://doi.org/10.5054/tq.2011.254523

Uchihara, T., & Clenton, J. (2020). Investigating the role of vocabulary size in second language speaking ability. Language Teaching Research, 24(4), 540–556. https://doi.org/10.1177/1362168818799371

Uchihara, T., & Clenton, J. (2022). The role of spoken vocabulary knowledge in second language speaking proficiency. *The Language Learning Journal. Advance online publication.* https://doi.org/10.1080/09571736.2022.2080856

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22*(2), 217–234. https://doi.org/10.1017/S0142716401002041

Webb, S. (2013). Depth of vocabulary knowledge. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics (pp. 1–8). Blackwell Publishing Ltd.*

Webb, S. A., & Sasao, Y. (2013). New directions in vocabulary testing. *Language Testing, 44*(3), 263–277. https://doi.org/10.1177%2F0033688213500582

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ILT International Journal of Applied Linguistics, 168(1), 33–69.* https://doi.org/10.1075/itl.168.1.02web

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *The Canadian Modern Language Review, 53*(1), 13–40. https://doi.org/10.3138/cmlr.53.1.13

Xing, P., & Fulcher, G. (2007). Reliability assessment for two versions of Vocabulary Levels Tests. *System, 35(2), 182–191.* https://doi.org/10.1016/j.system.2006.12.009

Zareva, A. (2005). Models of lexical knowledge assessment of second language learners of English at higher levels of language proficiency. *System, 33(4), 547–562.* https://doi.org/10.1016/j.system.2005.03.005

Zhang, B., & Annual, S. B. (2008). The role of vocabulary in reading comprehension: the case of secondary school students learning English in Singapore. *RELC Journal, 39(1), 51–76.* https://doi.org/10.1177%2F0033688208091140

Zhang, D. (2012). Vocabulary and grammatical knowledge in L2 reading comprehension: A structural equation modeling study. *Modern Language Journal, 96, 554–571.* https://doi.org/10.1111/j.1540-4781.2012.01398.x

Zhang, D., & Koda, K. (2017). Assessing L2 vocabulary depth with word associates format tests: Issues, findings, and suggestions. *Asian-Pacific Journal of Second and Foreign Language Education, 2(1), 1–30.* https://doi.org/10.1186/s40862-017-0024-0

Zhong, H. F. (2016). The relationship between receptive and productive vocabulary knowledge: a perspective from vocabulary use in sentence writing. *The Language Learning Journal, 46(4), 1–17.* https://doi.org/10.1080/09571736.2015.1127403

https://doi.org/10.17323/jle.2022.12901

The Interaction of Variables Affecting Definitional Skills: Extending Previous Research on Word Definitions

Zoe Gavriilidou, Angelos Markos, Chryssa Dourou

Democritus University of Thrace, Komotini, Greece

ABSTRACT

Background and Purpose. Given that definitional skills are closely related to literacy and reading comprehension, the purpose of this study was to extend the existing literature in definitional skills by empirically investigating the effect of new parameters that may affect word definitions and definitional types of content and form, such as grammatical categories, word structure, and semantic characteristics.

Methods. The sample consisted of 5152 recorded oral definitions produced by 322 individuals (pre-schoolers, school-age children, university students, and adults), who were asked to define 16 words orally. Definitions were transcribed and scored on a six-point scale along a continuum that reflects the developmental path of the definitions.

Results. The results indicated a significant interaction between grammatical category and word structure for content and form and also between word structure and semantic characteristics only for content. Furthermore, the grammatical category, word structure, and semantic characteristics were strongly associated with specific definitional types for content and form.

Implications. This paper broadens our knowledge on definitional skills and offers new insights into the variables that affect the production of definitions.

KEYWORDS

definitional skills, oral definition production, definition types, Greek Language

INTRODUCTION

Providing definitions, especially formal ones, is the ability to talk about word meanings which is strongly correlated to reading skills and school achievement (Anderson & Freebody, 1981; Snow Cancini, Gonzalez & Shriberg, 1989). It also guarantees depth in word knowledge and is a strong predictor of literacy skills, such as reading and writing, using books or the internet as a source of information, being familiar, producing and being able to find information in different genres and text types (descriptions, narrations, etc.) (Artuso, Palladino, Valentini & Belacchi, 2021, Snow et al., 1989). It may also play an active role in reading aloud (Nation & Cocksey, 2009).

However, special skills are required in order to formulate appropriate definitions even for words children or adults know well. Thus, these skills do not necessarily coincide with word knowledge.

The definitional skills have been investigated so far in the fields of education, psycholinguistics, and speech-language pathology. Previous research focused on the impact of literacy and school achievement (Artuso et al., 2021; Marinellie, 2010; Snow et al., 1989; Thorndike, Hagen, & Sattler, 1986; Watson, 1985), the effect of developmental characteristics or language impairment (Dosi & Gavriilidou, 2020; Dosi, Gavriilidou & Dourou, 2021; Gutierrez-Clellen & DeCurtis, 1999; Marinellie & Johnson, 2002) and finally the effect of first and second language (El Euch, 2007) on definitional skills.

Definitional skills also depend on the academic achievement, the verbal ability, and the intellectual performance of school-age children and adolescents

Citation: Gavriilidou, Z., Markos, A., & Dourou, C. (2022). The Interaction of Variables Affecting Definitional Skills: Extending Previous Research on Word Definitions. *Journal of Language and Education, 8*(3), 65-80. https://doi.org/10.17323/jle.2022.12901

Correspondence:
Zoe Gavriilidou
zoegab@otenet.gr

Received: August 19, 2021
Accepted: September 17, 2022
Published: September 30, 2022

(Nippold, 1995; Snow et al., 1989; Thorndike et al., 1986; Watson, 1985). On the other hand, formal word definitions demand an understanding of words, a vocabulary knowledge and, also, the ability to express that knowledge explicitly.

However, little is known about the effect of grammatical category, semantic characteristics, and word structure of words to be defined on definitional abilities and the types of definitions produced. Understanding the reasons why other children succeed and other fail in providing definitions of different types of words and how this affects their literacy and school achievement is crucial for preventing school failure. To brigde this gap, this study was designed to investigate how definitional abilities develop with respect to the above-mentioned variables.

LITERATURE REVIEW

Definitional Types of Content and Form

Definitions are combinations of informative contents and appropriate forms (Benelli, Belacchi, Gini & Lucangeli, 2006). More precisely, the canonical definitions should accomplish four formal (paraphrase, periphrastic form, phrasal autonomy and correct morpho-syntactic structure, and one content-related (semantic equivalence) requirements (Artuso et al., 2021).

The content of definitions refers to the strategies speakers employ in order to explain the meaning of a word; in other words, the semantics of definitions (Benelli, Arcuri & Marchesini 1988; Dourou, Gavriilidou & Markos, 2020; Gavriilidou, 2015; Marinellie & Johnson, 2002, 2004). These strategies include, among others, the following: (a) Functional definitions, which are a kind of informal definitions that build on the characteristics or properties that define the concept and describe what the item does or is used for (cutlery: we eat with it); (b) Descriptive definitions, which are definitions that describe the concept's perceptual appearance as in bicycle: it has a steering wheel, a saddle and two wheels; (c) Present state definitions, in which a person uses a situation to give a general reference to a particular place at a particular time, or to refer to what is happening to them as, for instance, in a question: what you are doing now; (d) Definition by Example, where different types of the concept are taken as examples, showing awareness of the aim of the message as directed to a receiver as in open-close: for example, I open the door and I close it; (e) Association/Result/Action. In this case the concept is associated with an event, a person, an experience or an action, or may mention the results that this concept brings (cheese-pie: my grandma's cheesepie); (f) Tautologic definitions, are erroneous definitions where the speakers repeat the same word twice (apple: it is an apple), or in case of compounds, they parse the concept in its components (sweet-sour: something that is sweet and sour); (g) Definition by self-reference is another strategy used for defining words. It refers to the ability of a person to speak of or refer to themselves, that is, to have the kind of thought expressed by the first person nominative singular pronoun "I" (clever: I am clever); (h) Class nonspecific definitions, where the concept is included in a wider category or related with the main elements but not with specific details (apple: a fruit); (i) Class specific definitions where a superordinate term to the concept is used (bicycle: it is a means of transport) or defining words by Synonyms is preferred (clever: intelligent); (j) Definitions by Combination are also possible. In this case speakers use a wider or a general category adding some extra characteristics (descriptive, functional etc.) of the concept (cheese-pie: an unhealthy food); (k) Aristotelian or Formal definitions are the most complete definitions where speakers use a superordinate term and the differentiating characteristics of the concept (question: it's a phrase that asks for answers).

Previous research has demonstrated that, with age, the content of definitions of words progresses from functional and concrete to more abstract and conventional (Anglin, 1977; Benelli et al., 1988; Dourou et al., 2020; Gavriilidou, 2015; Nippold, Hegel, Sohlberg & Schwarz, 1999; Watson, 1985;). No previous research has investigated so far how definitional types for content change according to the grammatical categories, semantic characteristics, and word structure of the words to be defined. However, the form of definitions refers to the syntactic patterns speakers use in order to shape the meaning of a word. Five (5) are the main definitional types regarding form (Dourou, et al., 2020; Gavriilidou, 2015; Marinellie, 2010; Marinellie & Johnson, 2002, 2004): (a) Nonverbal definitions, which are definitions where the speaker does not define the concept because it is unknown to him. There are no verbal responses, and gestures or tautologies may be used (apple: they show it); (b) Single word or Article + Word. In this case, speakers use only one word to define the concept with or without an article (cutlery: kitchen). Sometimes, speakers use only one simple sentence to define the concept, without any dependent clauses (cheese-pie: we eat it) (Definition with a Phrase, Clause or Simple Sentence); (c) Definitions with Transitional form where they use generic terms as "something" or "a thing" plus a modifying clause to define the concept (cheese-pie: something that has cheese inside); (d) Partially formal definitions where a superordinate term is used; however, without a complete syntactic form (question: it's a phrase with a question mark); (e) Formal definitions where speakers use a superordinate term and the differentiating characteristics of the concept (question: it's a phrase that asks for answers or information). In a more recent research, Belacchi & Benelli (2017) forwarded the idea of a definitional scale including five consecutive definitional levels; each level adds a new morpho-syntactic property, from single word definitions to the prototypical Aristotelian form, which ensures the appropriate expression of semantic contents.

Similarly to definitional types for content, types for form develop from the late preschool to the early school-age years. With increasing age, definitions develop from simple syntactic structure into Aristotelian form (Kurland & Snow, 1997). The results of previous research agree with more recent studies. More specifically, the category 'Phrase/Simple Clause' is the most frequent type of definition for preschoolers and elementary students, while for junior and senior high school students, university students and adults with university education (Dourou, 2019) the most preferred type is the 'Partially Aristotelian form'. According to the study of Marinellie and Johnson (2002), the use of Transitional form increases until 9-10 years old and then decreases. The most recent study by Dourou et al. (2020) found that the most common type of definition in form is the 'Phrase/Simple Clause', followed by the 'Partial Aristotelian definition'. The categories 'Transitional Form' and 'One Word or Article + Word' are placed in the mid-preferences while the Aristotelian form is the least preferred type. However, there is a gap in literature concerning how definitional types for content and form may vary depending on the grammatical categories, semantic characteristics, and word structure of the words to be defined.

The effect of grammatical categories on word definitions and definitional types

Grammatical category affects both the definitional ability and the definitional types chosen by the speakers. Previous research has shown that nouns receive higher scores in definitional tasks compared to verbs and adjectives. Hutenlocher and Lui (1979), respectively, insisted that the use of superordinate terms is less clear in verbs than in nouns. In the same line, Gentner (1982) and Miller (1991) suggested that verbs are difficult to define because they refer to activities, motion, changes of state, relations and all these reasons have an effect on the complexity of verbal definition production, while according to Graesser, Hopkinson & Schmid (1987), Gertner (1982) and Markman (1989) adjectives and verbs are dependent on nouns. Thus, definition skills for verbs and adjectives are less predictable and develop much later than noun definition skills (Johnson & Anglin 1995). Markowitz and Franz (1988) found that verb and adjective definitions are more variable in form than noun definitions, but verbs may have a conventional or a typical definitional form similar to nouns (migrate means moving from one place to another depending on the seasons). Johnson and Anglin (1995) examined the ability of elementary school children (aged 6 to 8) to define verbs and they came to the conclusion that verbs, compared to nouns, were more difficult to define because it was difficult to find a superordinate term, possibly due to the non-hierarchical structure of children's mental vocabulary. Marinellie and Johnson (2004) asked 30 Elementary students to define 10 nouns and 10 verbs and their findings showed no significant difference between nominal and verbal definitions with respect to content. On the other hand, their results revealed that form scores for nouns were significantly higher than those for verbs. Gavriilidou (2015) asked fifty-two (52) preschoolers to define sixteen (16) words (nouns, verbs and adjectives). The study found evidence that preschoolers had higher scores for nouns than for adjectives and verbs, while adjectives received higher scores compared to abstract nouns and verbs. To account for such differences, the author maintained that this may happen because, contrary to verbs which express change of state, mode, action, purpose, causality and usually lack a hierarchical structure, nouns are referential anchors, fact that facilitates the use of a hyperonymic term useful for their definition. Finally, Dourou (2019) examined the definitional ability of different age groups (preschoolers, lower vs. upper elementary students, Junior vs. Senior High school students and Low educated vs. Highly educated adults) and found evidence that all the above groups provided better definitions for nouns than for verbs and adjectives.

As stated before, the grammatical category also affects the definitional types provided during the definitional tasks. The study by McGhee-Bidlack (1991), who studied the way that adults define nouns, claimed - contrary to the findings of other research on children's definitions - that the majority of adults used superordinate terms for nominal definitions. In addition, most adults' definitions of nouns agree with the conventional form (an X is a Y that Z). Benelli et al. (2006) examined a sample of 280 children aged 6 to 12 and she found that nouns are mostly defined by introducing superordinate terms, whereas adjectives and verbs can also be defined by introducing synonyms. Marinellie and Chan (2006) observed that definitions of verbs produced by children at the age of 4 often include relationships and associations (*find*: you are happy when you find something), synonyms (*leave*: go) and verbs with a broad meaning in intransitive phrases (circulate: to move). Gavriilidou (2015) showed that preschoolers tend to give more functional definitions related to a particular event, person or place. The different types of definitions depend on the category to which a noun belongs (concrete/common/local/abstract nouns). The study also found that the most common definitional types for verbs were functional and definitions by example, while for adjectival definitions the most common types were descriptive and functional. As the author claimed, this finding can be justified because adjectives refer to the aesthetic properties of concrete objects or persons. Gandia (2016), showed that elementary school students expressed semantic content of nominal definitions through a synonym, descriptive characteristics, or through the function. Dourou (2019), examined the definitional ability of 50 Junior High school students and claimed that they had shown high performance in defining nominal definitions, both in content and form. The majority of the study participants used a combination of definitions and the class specific category, while the Partial Aristotelian form and the Aristotelian form were the most common types of their productions, with repect to syntax. The study also found that

the most common definitional types for verbs by Junior High school students was Tautology and Synonym, and the Phrase/Simple Clause definition with regard to form. In the same study, a majority of Senior High school students tended to prefer the Tautology for defining verbs and Phrase/Simple Clause definition, with regard to form. Dourou (2019) also concluded that highly educated adults preferred to define verbs using a class specific term and characteristics of the concept in content, while low educated adults showed a clear preference for Association/Result/Action. similarly, in form, highly educated adults preferred to define verbs with Partial Aristotelian form, while low educated adults with Phrase/Simple Clause.

Even though the effect of grammatical category on the production of definitions has been extensively studied, no previous research investigated the possible interaction of grammatical categories with other variables when speakers produce definitions nor how this interaction affects the definitional types chosen by the speakers.

The Effect of Morphological Structure on Word Definitions and Definitional Types

Simple/Derived and Compound Words

Words in Greek can be simple (e.g. *milo 'apple'*), derived (e.g. *xorevo 'dance'*), or compound (e.g. *iʎovasilema 'sunrise'*). According to the typology of Ralli (2005; 2013), there are four main categories of compounds: a) Stem+linking vowel+stem+inflection (e.g. kian+o+kran+os 'blue helmet'), b) stem+linking vowel+word (e.g. melan+o+doxio 'inkpot'), c) Word+stem+inflection (e.g. kato+sendon+o 'undersheet'), d) Word+word (e.g. ksana+pezo 'replay'). The two first categories are more productive while the other two are minor and usually less productive.

Greek also favors derivation (through affixation), mainly with suffixes to form nouns (xoreftis 'dancer', kalosini 'goodness'), verbs (xorevo 'dance', skoupizo 'wipe') and adjectives (melodikos 'melodic', melanxolikos 'melancholic') but also prefixes to form nouns (kataθesi 'deposition'), verbs (anavalo 'postpone') or adjectives (aoratos 'invisible').

Compounding is "one of the richest sources of word formation in everyday language and scientific terminology" (Ralli, 2005). Simple compounds are acquired early by children with typical development (Nicoladis, 2006). Compound acquisition seems to play a major role in vocabulary development. Children primarily treat compound words as single words, and then gradually become aware that they consist of two parts connected with a meaning relation between them. Children's knowledge of the meaning of compounds starts to develop from the preschool years and is affected by a series of factors such as the size of the compound family or the relation between the head and the modifier (Krott & Nicoladis, 2005; Nicoladis & Krott, 2007).

When it comes to derivation, on the other hand, previous research showed that typically developing children start to acquire derivation at an older age compared to inflectional morphology (Clark, 1998), and that they start using it productively at a later age. Marshall and Van Der Lely (2007:72) explain that this happens because "derivational suffixes are more irregular and constrained, and the form-to-meaning is not always predictable". According to Clark (1998), the acquisition of derivational affixes depends on frequency, semantic opacity, allomorphy and the presence or not of irregularities.

As a matter of fact, to define derived or compound words "children must learn to identify regularities in the relations between forms and meanings" (Clark & Berman, 1984; 1987) but also make assumptions about the contribution of each word part to the whole meaning of the word. Following the acquisitional principles of Clark (1981) and Clark & Hecht (1982), children's definitions may be facilitated if the derived or compound words to be defined are characterized by semantic transparency (known elements with one-to-one matches of meaning and form), formal simplicity (the less a word form changes the simpler it is), conventionality (for certain meanings a conventional word formation device exists) and productivity (some word formation devices are more productive than others in specific languages).

Dourou (2019) showed that the participants in her study provided better definitions in simple than in compound words. This is justified because the words in her study belong to the basic vocabulary of the students and demonstrate higher frequency than non-basic words. Moreover, her sample used *Associations or Result / Action* for defining simple words (xorevo '*dance*': get tired) and *Tautologies* for compounds (aspromavros 'black and white': black and white). A significant effect of education level (highly educated adults) on the definition of simple/derived words was also observed. More specifically, adults with university education had higher scores in the definition of simple/derived words and compound words than adults without university education. This is also depicted in the types of definitions chosen by both groups. Adults with high education preferred to define simple/derived and compound words combining a superordinate term with functional and descriptive characteristics of words or using a synonym, while low educated adults preferred Associations or Result / Action for defining simple words and Tautologies for compounds.

No previous research has investigated how morphological structure interacts with other variables during word definition productions and how this interaction affects the definitional types chosen by the speakers.

The Effect of Semantic Characteristics on Word Definitions and Definitional Types

Concrete vs. Abstract Nouns

Not only the grammatical categories of words affect the content and form of definitions but also the level of abstraction of the noun to be defined (Dourou, 2019; Gandia, 2016; Johnson & Anglin, 1995; McGhee-Bidlack, 1991; Nippold et al., 1999; Sadoski, Kealy, Goetz & Paivio, 1997). Three age groups (ages 10, 14, and 18) had higher performances on concrete nouns than abstract nouns in the study of McGhee-Bidlack (1991). Whereas the definitions of concrete nouns were based mainly on their superordinate terms and characteristics, abstract nouns were defined in terms of their characteristics, with their category terms often excluded. A gradual improvement in abstract definitions was observed with age but even at the age of 18, definitions of abstract nouns were far less formal, due to lacking the appropriate superordinate term. The ability of 96 children (aged 6 to 10) to define concrete and abstract nouns was investigated by Johnson & Anglin (1995). They found that definitions of concrete nouns were more precise due to their superordinate and subordinate connections with other nouns. In line with the results of other studies, they concluded that the hyperonyms of abstract nouns (e.g., "feeling for the concept of love",) are language skills, which have not yet been developed in lower elementary school students. For concrete nouns, the inclusion of a superordinate term is a skill that develops after the age of 7, when students tend to use an IS A-structure (*apple* is a round and red fruit) in form (syntactic structure) of definitions.

Sadoski, Kealy, Goetz & Paivio (1997) asked graduate students to produce written definitions for concrete and abstract nouns using computers. The results concerned the quality of definitions as well as the use of strategies and showed that when participants gave definitions of concrete nouns, they started earlier and wrote longer and higher quality definitions. In contrast, definitions of abstract nouns included more words in order to convey the abstract meaning.

Nippold et al. (1999) asked students (12 to 23 years old) to produce definitions of low frequency abstract nouns (e.g., burden, humility) that were presented in a random order. Findings showed that in abstract concepts, adults also encountered difficulties. Although the responses improved with age, only 58% of the oldest group responses were awarded full credit. The researchers concluded that the definitions of abstract nouns cannot be complete and precise until late adolescence.

In Gavriilidou (2015), preschoolers had higher scores when they defined concrete nouns compared to abstract ones. Furthermore, the students provided class-specific (e.g. apple: a fruit) or functional definitions for concrete nouns and erroneous or descriptive definitions for abstract ones. The author concluded that it is very difficult for students in early childhood to define abstract concepts because they have not developed the appropriate skills yet. This ability is a process that develops gradually from pre-adolescent and adolescent years to adulthood. Dourou (2019) found a gradual improvement in the definitions of abstract nouns with age. More specifically, from the last grades of elementary school, students are better at definitions of abstract nouns because they start to develop definitional skills for abstract entities. According to the content of definitions, the most frequent type of concrete noun definition was *Class-specific category*, while the most frequent type of abstract noun definition was *Association/Result and Action*. In terms of form, both for concrete and abstract nouns, the largest percentage of the sample provided definitions with the *Partial Aristotelian form*. It should be noted that most studies that investigated so far how the abstract/concrete distinction affects the production of definitions only take into account nouns. However, more recent research (Belacchi & Benelli, 2017) also investigated the effect of concreteness / abstractness on verb and adjective definitions.

To our knowledge, no previous study has investigated the effect of semantic characteristics on both content and form of definitions from preschool age to adulthood nor the possible interaction with the grammatical category that a word belongs to. Considering the privileged association of nouns with punctual concepts and that of adjectives and verbs with relational concepts and also that relational concepts are closely related to abstractness (Strik Lievers, Bolognesi & Winter, 2021), in this study, we opted to investigate the effect of concreteness / abstractness only on definitions of nouns.

Aims and Hypotheses

Taking into consideration previous gaps in the literature as indicated in the review, the general purpose of the present study was to extend the existing knowledge on the production of definitions by investigating new parameters that may affect definitions such as the effect of grammatical categories, word structure and semantic characteristics on definitional skills and definition types of content and form.

The first aim was to investigate the effect of grammatical categories (nouns, verb, adjectives) on word definitions and definitional type, both in content and form. Based on previous literature (Gavriilidou, 2015; Marinellie & Johnson, 2003, 2004), we expected that the utterances children produce for nouns, verbs and adjectives would differ. Specifically, and consistently with the literature on the effects of the grammatical categories of words on definitional skills (Dourou, 2019; Gavriilidou, 2015; Johnson & Anglin, 1995; Markowitz & Franz, 1988; McGhee-Bidlack, 1991; Nippold et al., 1999), we assumed that nouns would be better defined, compared

to adjectives and verbs, and would contain more hyperonyms and more precise and formal semantic content, compared to verbal and adjectival definitions. This hypothesis is based on the assumption that children, during the school years, have more practice in defining nouns than verbs and adjectives. On the other hand, we assumed that verbs and adjectives would be more difficult due to their non-hierarchical structure and would not contain relative clauses, such as nouns. An analogous development was also expected in form of nominal definitions.

The second aim was to study the effects of word structure (simple words, derivatives, and compounds) on word definitions and definitional types, both in content and form, in order to make predictions about the most conventional word formation device in Greek by grammatical category.

The third aim was to examine the effect of semantic characteristics (concrete and abstract words) in content and form of definitions. Taking into consideration previous literature on the effect of concreteness and abstraction on definitional skills (Gavriilidou, 2015; McGhee-Bidlack, 1991; Nippold, et al., 1999), we expected that concrete words would be defined easier than abstract ones. We predicted that the participants would define concrete nouns using a superordinate term (Johnson & Anglin, 1995), both in content and form, while abstract nouns would be defined with association and phrase/simple clause (Nippold et al., 1999).

METHODS

Participants

The study sample comprised 5152 recorded oral definitions classified for content and form (see Table 1 and 2 respectively) produced by a non-random sample of 322 individuals (140 males (43%) and 182 females (57%) of different age groups, from preschoolers to adults.

Assessments and Measures

The definition task included sixteen (16) words, of which eight (8) were nouns, four (4) verbs and four (4) adjectives. Eight (8) of them were compounds (4 nouns, 2 adjectives and 2 verbs) all constructed following the most frequent pattern of compounding in Greek: Stem+linking vowel+stem+inflection. For a detailed presentation of the procedure of word selection for the definition task see Dourou, et al. (2020). Effort was made to include semantically transparent and formally simple derivatives and compounds. The sixteen (16) words that were chosen were checked for their frequency and are depicted in Table 2.

The questionnaire was administered orally by the third author to each participant individually. It was deemed necessary to orally administer the instrument since the preschoolers included in the sample had not yet developed writing ability. Furthermore, oral administration minimized the risk of copying a definition through the internet or from a dictionary (electronic or printed).

The eight (8) nouns were interspersed with the four (4) verbs and four (4) adjectives in random order, and each participant was randomly assigned to an order. Following the research protocol of Marinellie & Johnson (2002, 2004), for nouns the investigator asked the usual question employed by teachers in Greek schools to elicit nominal definitions: Ti ine X; 'What is X?'. As this is the common practice in classroom, children could understand that they were asked to define a word from the grammatical class of 'noun'. The use of a natural prompt for nouns would maximize the chance that children would interpret as nouns common words with multiple meanings. In accordance with Marinellie & Johnson (2004), for each verb or adjective, on the other hand, the investigator asked children the usual question for eliciting verbal or adjectival definitions in Greek schools: Ti simeni X; 'What does X mean?'. The use of this natural prompt maximized the chances that the child would interpret these

Table 1

Distribution of subjects across gender and age groups

Age groups	Gender		Mean	Age Range	N
	M	F	Age	[Min, Max]	
Preschoolers	16	17	5.08	[5.00-5.70]	33
Lower Elementary	21	15	7.40	[6.00-8.20]	36
Upper Elementary	16	19	10.82	[8.80-12.1]	35
Junior High students	17	18	13.19	[12.0-14.4]	35
Senior High students	23	27	16.01	[15.0-17.2]	50
University students	24	59	21.41	[19.0-24.6]	83
Adults	23	27	57.40	[32.0-68.4]	50
Total	140	182	18.75	[5.00-68.4]	322

Table 2

Definitional Task Grouped per Word Categories

	Items	Grammatical Category	Semantic characteristics	Mode of Construction
Words from Textbooks of the Modern Greek Language of Elementary School	erotisi 'question'	Noun	Abstract	Derivative
	taksiði 'journey'	Noun	Abstract	Simple
	iʌovasilema 'sunset'	Noun	Abstract	Compound
	'sunrise'			
	makrozoia 'longevity'	Noun	Abstract	Compound
	tiropita 'cheese pie'	Noun	Concrete	Compound
	maçeropiruno 'cutlery'	Noun	Concrete	Compound
	aspromavros 'blackand white'	Adjective		Compound
	ylikoksinos 'sweet and sour'	Adjective	N/A	Compound
	aniyoklino 'open and close'	Verb		Compound
	siyotrayuðo 'hum'	Verb		Compound
Words from Gavriilidou's research (2015)	milo 'apple'	Noun	Concrete	Simple
	poðilato 'bicycle'	Noun	Concrete	Simple
	eksipnos 'intelligent'	Adjective		Simple
	astios 'funny'	Adjective	N/A	Derivative
	ðjavazo 'read'	Verb		Simple
	xorevo 'dance'	Verb		Derivative

words as verbs. Data collection lasted two months (October-December 2018). The study was approved by the Ethics Committee of the Department of Greek of Democritus University of Thrace. As children are involved in the study written consent was obtained from the legal guardians.

Data Scoring

Even though, a number of studies have shown the usefulness of a single coding system that jointly considers form and content (Benelli et al., 2006; Belacchi & Benelli, 2017; Artuso, et al., 2022), since form and content do not always change together and in the same ways (Johnson & Anglin, 1995; Litowitz, 1977; Wehren, DeLisi & Arnold, 1981;), for the needs of the present study we opted to investigate them separately and for that reason we adopted the scoring methodology of Marinellie & Johnson (2002, 2004) (as adapted in Dourou, 2019 and Gavriilidou, 2015).

Content

Comparing to the scoring of Marinellie & Johnson, (2002), four additional types of definitions (Present state, Tautology, Self-reference, Lexicographic definition) were added to the classification of definitions which emerged from the responses of the participants. Examples of content scoring are displayed in Table 3. Definitions were scored on a six-point scale along a continuum, from a minimum of 0 to a maximum of 5, to be consistent with a developmental progression suggested in previous literature. Low-level responses were *Function*, *Description*, *Present state*, *Example*, *Association* and *Tautology*. Mid-level responses were *Relation*, *Class non-specific*, *Class specific* and *Synonym*. High-level responses included *Combination I* and *II*, *Lexicographic definition* and *Aristotelian definition*.

Form

Examples of form scoring are displayed in Table 4. Definitions were also scored on a six-point scale along a continuum, to be consistent with a developmental progression suggested by the literature on definition. This scoring scheme was used in a study of the definitional skill of school-age children with specific language impairment (Marinellie & Johnson, 2002). Form categories included: *Nonverbal*; *Single Word or Article + Word*; *Phrase*, *Clause*, or *Simple Sentence*; *Transitional*; *Partial Aristotelian*; and *Aristotelian*. The highest possible form score for any participant was 80 points (16 words per participant, with a maximum of 5 points per word).

Table 3
Scoring Scheme used for the Content Categories

Content category	Example	Score
Error	milo [apple : ice-cream]	0
Function	tiropita [cheese pie: you eat it]	1
Description	milo [apple: red and round]	1
Present state	erotisi [question: what you are asking right now]	1
Example	aniγoklino [open and close: for example, open and close the door]	1
Association /Result /Action	diavazo [read: history]	1
Tautology	γlikoksinos [sweet-sour: sweet and sour]	1
Relation-Self-reference	eksipnos [intelligent: that's me]	2
Class non-specific	poðilato [bicycle: a thing]	2
Class specific	milo [apple: fruit]	3
Synonym	eksipnos [intelligent: clever]	3
Combination I	milo [apple: a thing that is red and round]	4
Combination II	poðilato [bicycle: means of transport with a steering wheel, saddle and pedal]	5
Lexicographic definition	diavazo [read: look at the words and understand their meaning]	5
Aristotelian definition	erotisi [question: a clause that asks for answers]	5

Table 4
Scoring Scheme used for the Form Categories

Form Category	Example	Score
Nonverbal	Participant demonstrates use of object or points to object	0
Single Word or Article + Word	iʎovasilema [sunrise: evening]	1
Phrase, Clause or Simple Sentence	milo [apple: we eat it]	2
Transitional form (use of "something" or "thing" plus modifying clause)	erotisi [question: something that calls for answers]	3
Partial Aristotelian form	milo [apple: a fruit]	4
Aristotelian form	poðilato [bicycle: means of transport with a steering wheel, saddle, pedal and without motor]	5

Reliability

Content and Form

Interrater reliability of content coding was evaluated for all responses given by 64 participants. This refers to 20% of the data (1024 definitions). Identically coded responses were considered an agreement. The two raters were the first author and a PhD student of the Department of Greek Philology of the Democritus University of Thrace. The investigator's grade was blinded. The percentage of agreement was calculated by dividing the number of responses coded identically by the total number of coded responses (1024 definitions). For the content of the responses, 908 common responses were recorded indicating an inter-rater agreement of 88.6%.

Inter-rater reliability of form coding was evaluated for all responses in the same way as content coding. For the form of the responses, 952 common responses were recorded indicating an inter-rater agreement of 92.9%.

Data Analysis

To investigate the effect of grammatical categories (noun, adjective, verb) and word structure (simple, compound, derivative) on content and form of definitions (scores between 0 and 5), a two-way MANOVA was conducted, followed by univariate ANOVAs with Tukey's HSD post-hoc test. Since the distinction concrete/abstract is relevant for nouns only, to investigate the effect of semantic characteristics and word structure on content and form of noun definitions another

two-way MANOVA was conducted. The assumptions of normality, linearity, homogeneity of variances and covariances were met in both cases. MANOVAs were conducted with IBM SPSS Statistics 23.0.

Correspondence Analysis – CA (Greenacre, 2007) was employed to investigate the associations between grammatical categories, word structure and semantic characteristics, and the definitional types of form and content. Specifically, CA was applied to (a) the contingency table formed by the combined categories of grammatical, semantic and word structure (e.g., noun/abstract/derived, noun/abstract/simple, etc.) and the six definitional types of form and (b) the contingency table formed by the combined categories described in the previous point and thirteen out of the fifteen definitional types of content. Present state definitions and definitions by self-reference were omitted due to low number of cases (see Table 1). To display the CA results, the contribution biplot scaling was used to facilitate interpretation (Greenacre, 2013). In the contribution biplot, the combined categories are displayed as points on a two-dimensional map, where the distances between them are spatially interpretable. The definitional types are displayed as vectors or biplot axes on the same map. The length of a biplot axis indicates the importance or contribution of the corresponding definitional type to the solution; biplot axes lying close to (or far from) the origin contribute little (or considerable) to the solution. Moreover, the perpendicular projection of a point (combined category) on a biplot axis (definitional type), approximates the frequency of the corresponding definitional type for that category. CA was applied with the package factoextra (Kassambara & Mundt, 2020) in R 4.0.0.

RESULTS

The Effect of Grammatical Categories and Word Structure on Content and Form of Definitions

The results of two-way MANOVA, with two independent variables – grammatical categories and word structure – and two dependent variables – content and form of definition scores, indicated a statistically significant interaction effect between grammatical category and word structure on the combined dependent variables, $F(6, 10286) = 54.746$, $p < .001$, Wilks' $\Lambda = .939$, partial $\eta 2 = .031$. Follow up univariate two-way ANOVAs showed a statistically significant interaction effect between grammatical categories and word structure for both content, $F(3, 5144) = 64.576$, $p < .001$, partial $\eta 2 = .036$, and form scores, $F(3, 5144) = 30.613$, $p < .001$, partial $\eta 2 = .018$. As such, simple main effects analyses were conducted for both types of scores.

With regard to content definition scores, there were statistically significant differences between word structure categories for nouns, $F(2, 5144) = 39.276$, $p < .001$, partial $\eta 2 = .015$, verbs, $F(2, 5144) = 46.397$, $p < .001$, partial $\eta 2 = .018$ and adjectives, $F(1, 5144) = 77.15$, $p < .001$, partial $\eta 2 = .015$. The mean content definition scores for simple, compound and derived nouns were 3.09, 2.98 and 2.18, respectively. Simple nouns were significantly better defined than derived nouns, .91, 95% CI [.66, 1.16], $p < .001$ and compound nouns were significantly better defined than derived nouns, .80, 95% CI [.55, 1.04], $p < .001$. The mean difference between simple and compound nouns was not statistically significant, .11, 95% CI [–.05, .28], $p = .32$. The mean content definition scores for simple, compound and derived verbs were 1.95, 1.68 and 2.74, respectively. Derived verbs were significantly better defined than simple verbs, .80, 95% CI [.40, 1.10], $p < .001$ and compound verbs, 1.06, 95% CI [80, 1.33], $p < .001$. The mean difference between simple and compound verbs was not statistically significant, .27, 95% CI [–.02, .47], $p = .06$. Last, the mean content definition scores for simple and compound adjectives were 2.55 and 1.76, respectively. Simple adjectives were significantly better defined than compound adjectives, .80, 95% CI [.62, .97], $p < .001$. The estimated marginal means and confidence intervals of content definition scores for grammatical categories and word structure are shown in Figure 1.

With regard to form definition scores, there were statistically significant differences between word structure categories for nouns, $F(2, 5144) = 41.716$, $p < .001$, partial $\eta 2 = .016$, verbs, $F(2, 5144) = 29.527$, $p < .001$, partial $\eta 2 = .011$, but not for adjectives, $F(1, 5144) = .63$, $p = .427$. The mean form definition scores for simple, compound and derived nouns were 3.42, 3.01 and 2.78, respectively. Simple nouns were significantly better defined than compound nouns, .34, 95% CI [.22, .46], $p < .001$ and derived nouns, .64, 95% CI [.45, .82], $p < .001$, and compound nouns were significantly better defined than derived nouns, .29, 95% CI [.12, .47], $p < .001$. The mean form definition scores for simple, compound and derived verbs were 2.73, 2.27 and 2.83, respectively. Simple and derived verbs received significantly better scores than compound verbs, .46, 95% CI [.26, .65], $p < .001$ and .56, 95% CI [.36, .75], $p < .001$, respectively. The mean difference between simple and derived verbs was not statistically significant, .10, 95% CI [-32, .13], $p = .82$. The estimated marginal means and confidence intervals of form definition scores for grammatical categories and word structure are shown in Figure 2.

The Effect of Semantic Characteristics and Word Structure on Content and Form of Definitions

A second two-way MANOVA was run with word structure and semantic characteristics as independent variables and content and form of noun definition scores as dependent variables. There was a statistically significant interaction effect between grammatical category and word structure on

Figure 1

Estimated Marginal Means of Content Score by Grammatical Category and Word Structure. Error Bars represent 95% Confidence Intervals.

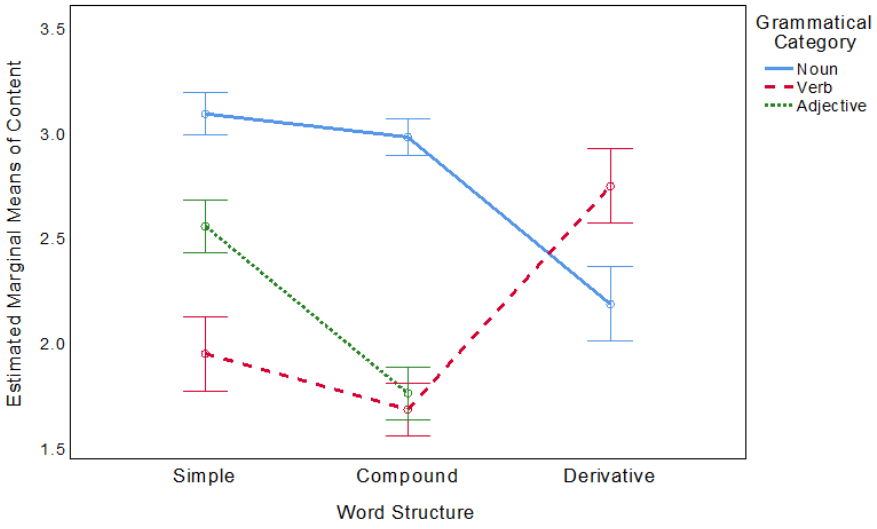
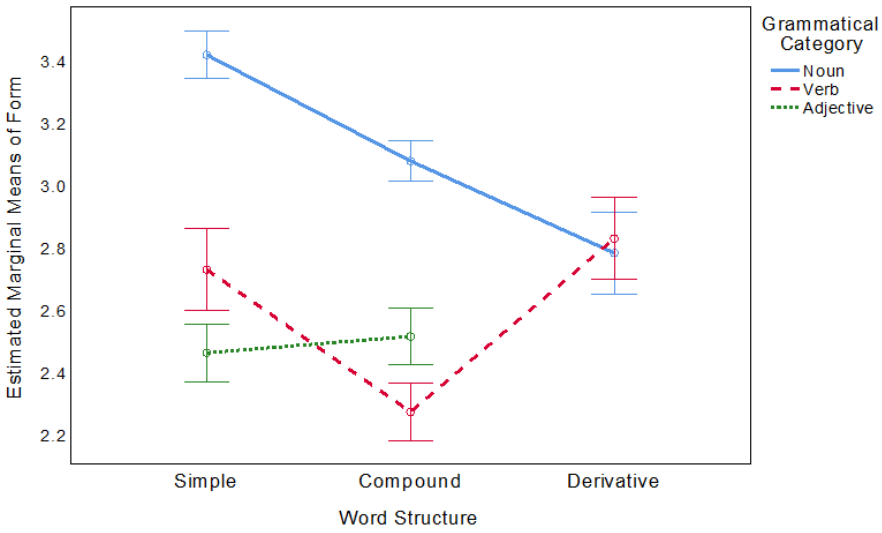


Figure 2

Estimated Marginal Means of Form Score by Grammatical Category and Word Structure. Error Bars Represent 95% Confidence Intervals.



the combined dependent variables, $F(2, 2570) = 25.164$, $p < .001$, Wilks' $\Lambda = .981$, partial $\eta 2 = .019$. Follow up univariate two-way ANOVAs showed a statistically significant interaction effect between grammatical categories and word structure for content scores, $F(1, 2571) = 87.566$, $p < .001$, partial $\eta 2 = .012$, but not for form scores, $F(1, 2571) = .261$, $p < .001$. As such, a simple main effects analysis was conducted for content scores only.

With regard to content definition scores, there were statistically significant differences between semantic characteristics for simple nouns, $F(1, 2571) = 43.071$, $p < .001$, partial $\eta 2 = .016$, but not for compound nouns, $F(1, 2571) = .729$, $p = .393$ (note that there were no concrete derived nouns). The mean content definition scores for simple concrete and simple abstract nouns were 3.34 and 2.59, respectively. Simple concrete nouns were significantly better defined than simple abstract nouns, .74, 95% CI [.52, .97], $p < .001$. Abstract

derived nouns were the most difficult category of nouns to define. The estimated marginal means and confidence intervals of content definition scores by word structure and semantic categories are shown in Figure 3.

With regard to form definition scores, the main effects of semantic characteristics, $F(1, 2571) = 174.251$, $p < .001$, partial $\eta 2 = .063$ and word structure, $F(2, 2571) = 7.264$, $p = .001$, partial $\eta 2 = .006$, were statistically significant. The mean form scores for concrete and abstract nouns were 3.55 and 2.80, respectively. Concrete nouns were significantly better defined than abstract nouns, .75, 95% CI [.65, .85], $p < .001$, independently of the word structure. The estimated marginal means and confidence intervals of form definition scores by word structure and semantic categories are shown in Figure 4.

Associations between Grammatical Categories, Word Structure, and Semantic Characteristics with Types of Definitions

Figures 5a and 5b show the main results of Correspondence Analysis on the contribution biplots of the first and second principal axes. These maps allow us to examine and reveal the associations between grammatical categories, word structure, and semantic characteristics (as combined categories) and the definitional types of form and content.

With regard to form, the first principal axis explained 52.6% of the total inertia (or variance in the data) and the second principal axis explained 36.1% (hence 88.7% in total); none of the remaining three principal axes explained more than 7%. This suggests that a two-dimensional CA solution gives a good approximation of the data. The origin of the map (0,0) represents the average profile of definition preference. Along the first principal axis, simple and compound concrete nouns (NounConcSimp, NounConcComp) tend to be associated with the Partially Aristotelian form and the Aristotelian form (ParArstl5, Arstl6). On the other hand, these are the least common types of definitions for abstract derived nouns (NounAbsDeriv) and compound adjectives (AdjComp). Along the second principal axis, simple adjectives (AdjSimple) are strongly associated with the Transitional form (Trans4), whereas compound verbs (VerbComp) are strongly associated with definitions with a Phrase, Clause or Simple Sentence (PhrsClsSimSent3).

With regard to content, the first principal axis explained 35% of the total inertia and the second, third and fourth axes explained 29.1%, 14.5% and 11.2%, respectively (hence 90% in total). None of the remaining five principal axes explained more than 6%, suggesting that a four-dimensional CA solution best explains the data. In Figure 5b, we plot the first two principal axes, but the third and fourth axes were also inspected. Along the first principal axis compound verbs and compound adjectives (VerbComp, AdjComp) tend to be associated with Tautologic (Tau7) and Example (Exmpl5) definitions. Along the second axis, simple concrete nouns (NounConcSimp) are associated with class specific definitions (Classpc10) and simple verbs and adjectives (VerbSimp, AdjSimp) are associated with Synonym and Association /Result /Action definitions (Syn11, AscResAct6). Along the third axis, concrete compound nouns (NounConcComp) are associated with Class non-specific and Function definitions (Classnn9, Fun2). Last, along the fourth axis, abstract compound nouns (NounAbsComp) are associated with Combination II and Lexicographic definitions (CombII13, LexDef14).

Figure 3

Estimated Marginal Means of Content Score by Word Structure and Semantic Characteristics. Error Bars represent 95% Confidence Intervals.

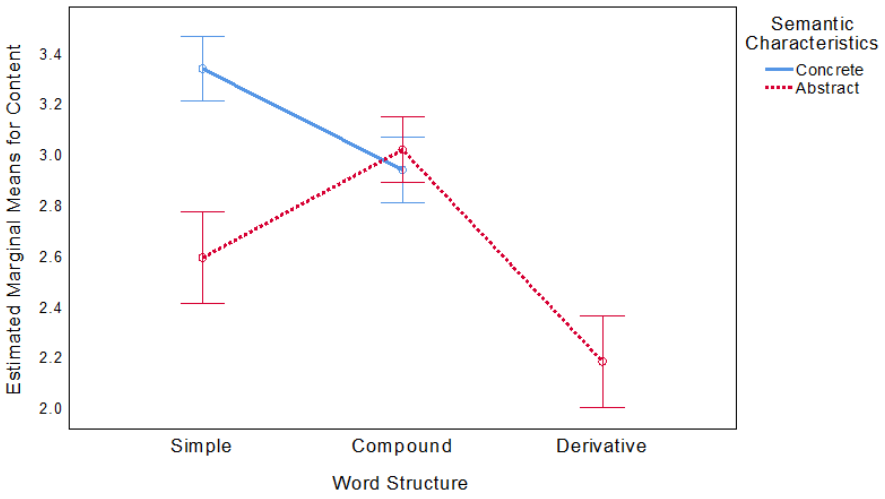


Figure 4

Estimated Marginal Means of Form Score by Word Structure and Semantic Characteristics. Error Bars Represent 95% Confidence Intervals.

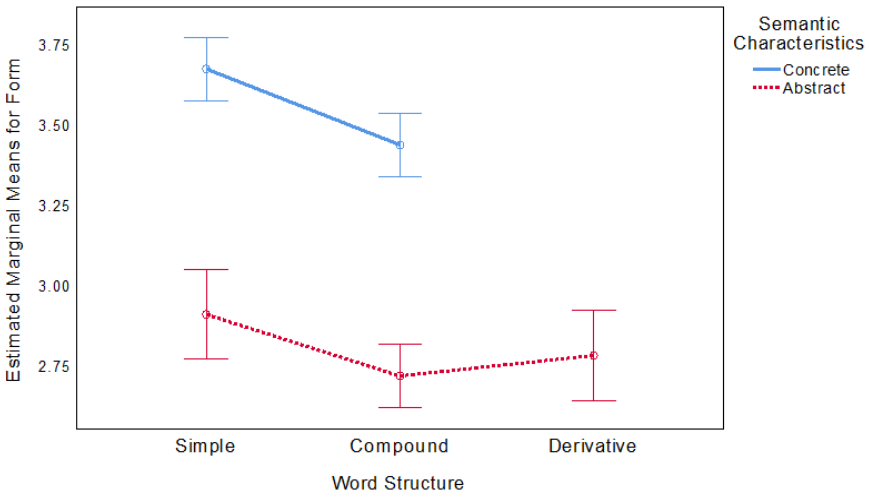
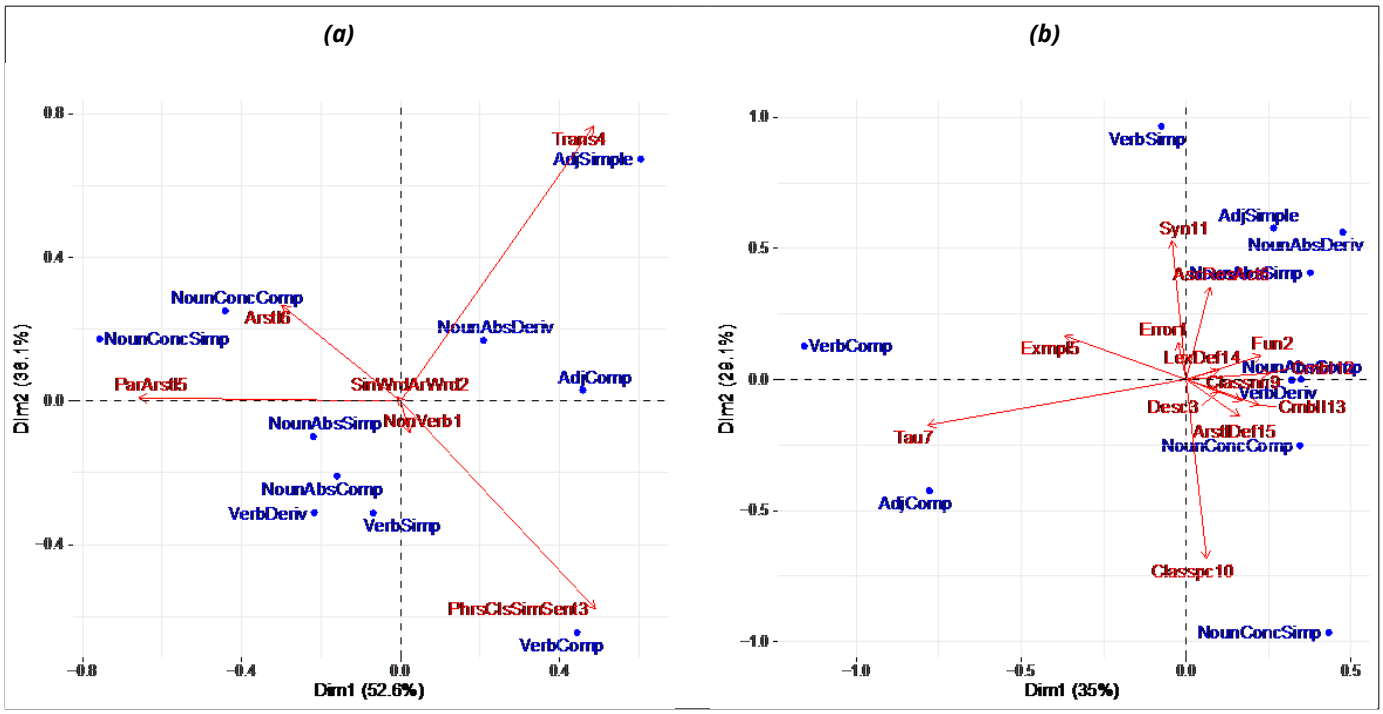


Figure 5

Contribution Biplots of CA on Grammatical Categories, Word Structure and Semantic Characteristics (as Combined Categories) and the Definitional Types of Form and Content.



DISCUSSION

The purpose of the present paper was to investigate the effect of grammatical categories (nouns, verbs and adjectives), word structure (simple words/derivatives vs. compounds), and semantic characteristics (concrete vs. abstract) on word definitions and definitional types of content and form.

Our first aim was to check the effect of grammatical categories on content and form of the produced definitions. It was expected that nouns would have higher scores compared to adjectives and verbs. This study verified our hypothesis. Confirming previous research (Dourou, 2019; Gavriilidou, 2015; Johnson & Anglin, 1995), the results showed that simple and compound nouns received better scores in defini-

tions, followed by simple and compound adjectives and, finally, simple and compound verbs. However, interestingly, a statistically significant interaction between grammatical category and word structure, not only for content but also for form, was found in our data that had not been mentioned in previous research: derived verbs were better defined than derived nouns suggesting that, in this case, it is not the distributional property of noun vs. verb per se, but the interaction between this property and word structure that had a greater influence. Future research should offer cross linguistic data for answering the question of how these two parameters interact.

Our second aim was to investigate the effect of word structure (simple words, derivatives and compounds) on word definitions and definitional types, both on content and form. Our findings contradict those reported in Dourou (2019) who found better performance of her sample in simple than in compound words. As discussed in the previous paragraph, a statistically significant interaction between grammatical category and word structure was found in our research. Our results showed that simple or compound nouns were better defined than derived nouns, while the opposite was the case with verbs and adjectives, where derived verbs received better scores than simple and compound verbs and derived nouns. In other words, the sample was better able to provide word definitions for simple or compound nouns than derived nouns but also defined better derived verbs compared with simple and compound verbs or derived nouns. This result may indicate that children find more transparent compounding compared to derivation for nouns and derivation compared to compounding for verbs and this may reflect that, in Greek, compounding is the more conventional and productive word formation device for nouns, while derivation is the most conventional word formation device for verbs. This finding should be verified with data from lexical statistics describing Greek vocabulary. Overall, this result indicates that the performance in definitions is a rather complex phenomenon and we should not look at word structure differences in children's definitions in a single, oversimplified way, but rather account for them in relation to grammatical category, semantic characteristics, or other variables.

Our third aim was to examine the effect of semantic characteristics (concrete and abstract words) on content and form of definitions. It was expected to find better scores for concrete nouns than for the abstract ones both in content and in form. The results of the present study partially confirmed our hypothesis, since statistically significant differences were found only for content suggesting that semantic characteristics of the word to be defined have an impact mainly on the information included in the definition, in other words the semantics of it. An interesting statistically significant interaction between semantic characteristics and word structure was also found; simple concrete nouns were better defined than simple abstract nouns in content. The results also showed that abstract derived nouns are the most difficult category to be defined. Thus, the findings of the present study extend previous work (Gavriilidou, 2015; Johnson & Anglin, 1995; Nippold et al.) by showing that semantic characteristics interrelate with other variables such as word structure.

Finally, with respect to the effect of grammatical categories, word structure and semantic characteristics on definition types of content and form provided by our sample it was found that, for form, simple and compound concrete nouns tend to be associated with the Partially Aristotelian form and the Aristotelian form. This finding extends previous studies (Markowitz & Franz, 1988; Snow, 1990) which found that nouns are defined with Class specific, Aristotelian or formal definitions, including a superordinate term plus distinguishing characteristics in a modifying clause, and suggests that the type of definitions depends not on isolated variables but on an interaction of parameters and that categories that are easier to be defined associate mainly with the Aristotelian form in definitions. With regard to content, it was found that: (1) compound verbs and compound adjectives tend to be associated with Tautologic definitions; This finding may be accounted for by the fact that, in their effort to define a compound word, speakers tend to analyse its components, thus often arriving to tautologic definitions; (2) simple concrete nouns are associated with class specific definitions, suggesting that concrete nouns may activate more easily hyperonymic terms in their definition than other categories of words; these terms function as an anchor around which the definition is formed; (3) simple verbs and adjectives are associated with definitions by a Synonym, reflecting teachers' classroom practices that systematically promote synonym definitions for verbs and adjectives; (4) concrete compound nouns are associated with Class non-specific definitions.

These findings also extend previous studies (Marinellie & Chan, 2006; Marinellie & Johnson, 2003).

Limitations

There are two major limitations of this study. First, following the research protocol of Marinellie & Johnson (2004), definitions for nouns were elicited through a different prompt than definitions for verbs and adjectives. This may have had an impact on the results. Second, only one derived adjective was included in the study.

Despite any limitations, it is hoped that this study provides additional understanding of how speakers define words and that this research-based knowledge can be employed in helping secondary or university teachers to expand their perspectives on good teaching to their pupils/students of how to effectively define different types of words.

CONCLUSION

To summarize, this study investigated the effect of grammatical categories, word structure, and semantic characteristics on word definitions and definitional types of content and form. Its significance lies in the fact that it has extended research on definitional skills and definitional types by adding new dimensions of variation in definitional skills, such as word structure, and extending prior knowledge of dimensions like grammatical category or semantic characteristics.

An interaction between grammatical category, semantic category and word structure was found providing evidence about the theoretical assumption that the performance in definitions is a complex phenomenon and we should not look at isolated variables in an oversimplified way, but instead we should investigate the interaction of multiple parameters. The present study also offered useful insights about the most conventional word formation device in Greek by grammatical category; compounding for nouns and derivation for verbs this finding, however, needs to be verified with further research.

Future investigation should test with Greek data how effective a single coding system that jointly considers form and content would be. Furthermore, it would be interesting to use Greek data to check the effect of concreteness/ abstractness on definitions for verbs and adjectives.

ACKNOWLEDGMENTS

This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Program «Human Resources Development, Education and Lifelong Learning» in the context of the project "Reinforcement of Postdoctoral Researchers - 2nd Cycle" (MIS-5033021), implemented by the State Scholarships Foundation (IKY).

DECLARATION OF COMPETING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Zoe Gavriilidou: conceptualization, methodology, project administration, resources, supervision, validation, writing-review & editing.

Angelos Markos: conceptualization, data curation, investigation, methodology, project administration, validation, visualization, writing-original draft, formal analysis.

Chryssa Dourou: conceptualization, formal analysis, methodology, project administration, supervision, validation, visualization, writing-original draft, writing-review & editing.

REFERENCES

Anderson,R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 97-117). International Reading Association.

Anglin, J. (1977). *Word, object and conceptual development*. Norton.

Artuso, C., Palladino, P., Valentini, P. & Belacchi, C. (2021). Definitional skills as a bridge towards school achievement. *Sustainability*, *14*(1). https://doi.org/10.3390/su14010286

Belacchi, C., & Benelli, B. (2017). "A mosaic is not a place where you go to see pictures': On the nature of definitions and the development of definitional skills. *Rivista Di Psicolinguistica Applicata*, *17*(1), 71–91.

Benelli, B., Arcuri, L., & Marchesini, G. (1988). Cognitive and linguistic factors in the development of word definitions. *Journal of Child Language,* 15(3), 619-635. https://doi.org/10.1017/S0305000900012599

Benelli, B., Belacchi C., Gini, G., & Lucangeli, D. (2006). To define means to say what you know about thing: The development of definitional skills as metalinguistic acquisition. *Child Language, 33*(1), 71-97. https://doi.org/10.1017/S0305000905007312

Clark, E. V. (1981). Lexical innovations: How children learn to create new words. In W. Deutsch (Ed.), *The child's construction of language* (pp. 299-328). Academic Press.

Clark, E. V. (1998). The acquisition of morphology. In A. J. Spencer & A. M. Zwicky (Eds.), *Handbook of morphology* (pp. 374-389). Blackwell.

Clark, E. V., & Berman, R. A. (1984). Structure and use in the acquisition of word formation. *Language*, *60*(3), 542-590.

Clark, E. V. & Berman, R. A. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child Language, 14*, 547–67. https://doi.org/10.1017/S030500090001028X

Clark, E. V., & Hecht, B. F. (1982). Learning to coin agent and instrument nouns. *Cognition*, *12*(1), 1–24. https://doi.org/10.1016/0010-0277(82)90027-0

Dosi, I. & Gavriilidou, Z. (2020). The role of cognitive abilities in the development of definitions by children with and without Developmental Language Disorder. *Journal of Psycholinguistic Research*, *49*, 761-777. https://doi.org/10.1007/s10936-020-09711-w

Dosi, I., Gavriilidou, Z. & Dourou, A. (2021). Definitional skills of learners with and without developmental language disorder. *International Journal of Learning, Teaching and Educational Research, 20*(10), 193-216. https://doi.org/10.26803/ijlter.20.10.11

Dourou, C. (2019). *The comparison of definition ability of different age groups* [Unpublished doctoral dissertation]. Democritus University of Thrace.

Dourou, C., Gavriilidou, Z. & Markos, A. (2020). Definitional skills and preferred definition types according to age, gender, educational level and career orientation. *International Journal of Research Studies in Education, 9*(2), 29-49. https://doi.org/10.5861/ijrse.2020.5021

El Euch, S. (2007). Concreteness and language effects in the quality of written definitions in L1, L2 and L3. *International Journal of Multilingualism, 4*(3), 198-216. https://doi.org/10.2167/ijm071.0

Gandia, A.M. (2016). *On word definition in children and adults: Effects of word category and level of abstraction* [Doctoral dissertation]. Universitat de Barcelona.

Gavriilidou, Z. (2015). The development of noun, verb and adjective definitional awareness in Greek preschoolers. *Journal of Applied Linguistics, 30*, 44-58.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity vs. natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Language, culture, and cognition* (pp. 301–335). Erlbaum.

Graesser, A. C., Hopkinson, P. & Schmid, C. (1987). Differences in interconcept organization between nouns and verbs. *Journal of Memory and Language, 26*(2), 242-253. https://doi.org/10.1016/0749-596x(87)90126-4

Greenacre, M. (2007). *Correspondence analysis in practice*. CRC press.

Greenacre, M. (2013). Contribution biplots. *Journal of Computational and Graphical Statistics, 22*(1), 107-122.

Gutierrez-Clellen, V., & DeCurtis, L. (1999). Word definitional skills in Spanish speaking children with language impairment. *Communication Disorders Quarterly, 21*(1), 23-31. https://doi.org/10.1177/152574019902100104

Huttenlocher, J., & Lui, F. (1979). The semantic organization of some simple nouns and verbs. *Journal of Verbal Learning and Verbal Behavior, 18*, 141–162. https://doi.org/10.1016/S0022-5371(79)90091-4

Johnson, C. J., & Anglin, J. M. (1995). Qualitative developments in the content and form of children's definitions. *Journal of Speech and Hearing Research*, *38*, 612-629. https://doi.org/10.1044/jshr.3803.612

Kassambara, I., & Mundt, F. (2020). Factoextra: Extract and visualize the results of multivariate data analyses. Version 1.0.7. https://CRAN.R-project.org/package=factoextra

Krott, A. & Nicoladis, E. (2005). Large constituent families help children parse compounds. *Journal of Child Language*, *32*(1), 139–58. https://psycnet.apa.org/doi/10.1017/S0305000904006622

Kurland, B.F. & Snow, C. (1997). Longitudinal measurement of growth in definitional skill. *Journal of Child Language 24*, 603–26. https://doi.org/10.1017/s0305000997003243

Litowitz, B. (1977). Learning to make definitions. *Journal of Child Language*, *4*(2), 289–304. https://doi.org/10.1017/S0305000900001665

Marinellie, S. (2010). Improving children's formal word definitions: A feasibility study. *Child Language Teaching and Therapy*, *26*(1), 23-37. https://doi.org/10.1177/0265659009349970

Marinellie, S.A. & Chan Y. (2006). The effect of word frequency of noun and verb definitions: A developmental study. *Journal of Speech, Language, and Hearing Research, 49*(5), 1001–1021. http://dx.doi.org/10.1044/1092-4388(2006/072)

Marinellie, S. A., & Johnson, C. (2002). Definitional skill in school-age children with specific language impairment. *Journal of Communication Disorders*, *35*(3), 241-259. https://doi.org/10.1016/S0021-9924(02)00056-4

Marinellie, S. A., & Johnson, C. (2003). Adjective definitions and the influence of word frequency. *Journal of Speech, Language & Hearing Research*, *46*(5), 1061–76. https://doi.org/10.1044/1092-4388(2003/084)

Marinellie, S. A., & Johnson C. (2004). Nouns and verbs: A comparison of definitional style. *Journal of Psycholinguistic Research*, *33*(3), 217-235. https://doi.org/10.1023/B:JOPR.0000027963.80639.88

Markman, E.M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.

Markowitz, J., & Franz, S. (1988). The development of defining style. *International Journal of Lexicography, 1*, 253-267.

Marshall, C. R., & van der Lely, H. K. J. (2007). Derivational morphology in children with grammatical-specific language impairment. *Clinical Linguistics & Phonetics*, *21*, 71-91. https://doi.org/10.1080/02699200600594491

McGhee-Bidlack, B. (1991). The development of noun definitions: A metalinguistic analysis. *Journal of Child Language, 18*, 417-434. https://doi.org/10.1017/S0305000900011132

Miller, G. A. (1991). *The science of words*. Freeman.

Nation, K. & Cocksey, J. (2009). The relationship between knowing a word and reading it aloud in children's word reading development. *Journal of Experimental Child Psychology, 103*(3)*,* 296-308. https://psycnet.apa.org/doi/10.1016/j.jecp.2009.03.004

Nicoladis, E. (2006). Preschool children's acquisition of compounds. In G. Libben & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 96-124). Oxford University Press.

Nicoladis, E., & Krott, A. (2007). Family size and French-speaking children's segmentation of existing compounds. *Language Learning*, *57*(2), 201–28. https://doi.org/10.1111/j.1467-9922.2007.00407.x

Nippold. M. A. (1995). School-age children and adolescents norms for word definition. *Journal of Language, Speech, and Hearing Services in Schools*, *26*(4), 320-325. https://psycnet.apa.org/doi/10.1044/0161-1461.2604.320

Nippold, M. A., Hegel, S., Sohlberg, M. M., & Schwarz, I. (1999). Defining abstract entities: Development in preadolescents, adolescents, and young adults. *Journal of Speech, Language, and Hearing Research*, *42*, 473–481. https://doi.org/10.1044/jslhr.4202.473

Ralli, A. (2005). *Morphology.* Patakis.

Ralli, A. (2013). *Compounding in modern Greek.* Springer. https://doi.org/ 10.1007/978-94-007-4960-3

Sadoski, M., Kealy, W. A., Goetz, E. T., & Paivio, A. (1997). Concreteness and imagery effects in the written composition of definitions. *Journal of Educational Psychology*, *89*(3), 518-526. https://doi.org/doi/10.1037/0022-0663.89.3.518

Snow, C. E. (1990). The development of definitional skill. *Journal of Child Language*, *17*(3), 697-710. https://doi.org/10.1017/S0305000900010953

Snow, C. E., Cancini, H., Gonzalez, P., & Shriberg, E. (1989). Giving formal definitions: An oral language correlate of school literacy. In D. Bloome (Ed.), *Literacy in classrooms* (pp. 233–249). Ablex.

Strik Lievers, F., Bolognesi, M., & Winter, B. (2021). The linguistic dimensions of concrete and abstract concepts: lexical category, morphological structure, countability, and etymology. *Cognitive Linguistics*, *32*(4), 641-670. https://doi.org/10.1515/cog-2021-0007

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Technical manual: Stanford-Binet Intelligence Scale* (4th ed.). Riverside.

Watson, R. (1985). Towards a theory of definition. *Journal of Child Language*, *12*(1), 181-197. https://doi.org/10.1017/S0305000900006309

Wehren, A., De Lisi, R., & Arnold, M. (1981). The development of noun definition. *Journal of Child Language*, *8*, 165-175. https://doi.org/10.1017/S0305000900003081

https://doi.org/10.17323/jle.2022.13608

Digital Support for Teaching Punctuation in Academic Writing in English

Marina Ivanova [1], **Nadezhda Arupova** [2], **Natalya Mekeko** [1]

[1] RUDN University, Moscow, Russia

[2] MGIMO University, Moscow, Russia

ABSTRACT

Background. Academic writing skills are crucial to the enterprise of higher education, because much of the academic communication is in writing. As a rule, foreign language learners face different problems with vocabulary misuse, grammatical errors, spelling, capitalization, punctuation and some others when write academic texts in English. There are various technologies for solving these problems. One of them is digital support, because traditional types of academic writing instruction in the classroom are not always sufficient.

Purpose. The present study researches the influence of digital support on students' knowledge and punctuation skills in academic writing in English.

Methods. The paper summarizes the results of an empirical study: training punctuation in academic writing lessons for two groups of students. A total of 24 students aged 20 to 24 participated in the experiment. The control group was applied a face-to-face and a text-book based traditional methodology. The experimental group was trained not only with a basic text-book but also with digital support. The level of knowledge and abilities in punctuation were measured with three final tests.

Results. The results obtained indicate that academic writing instructions in traditional text-books are not enough to develop punctuation skills. In its turn, digital support in training punctuation has significant didactic potential: the experimental group has demonstrated higher results in the final tests than the control one. Thus, digital support should be one of the main parts of the academic writing training process and should be included in training programs.

Implications. The materials of the present paper (resources on punctuation, evaluating tables, types of tasks and tests) can be used in a real teaching practice.

KEYWORDS

high education, academic writing, digital support, writing skills, punctuation

Citation: Ivanova, M., Arupova, N., & Mekeko, N. (2022). Digital Support for Teaching Punctuation in Academic Writing in English. *Journal of Language and Education, 8*(3), 81-96. https://jle.hse.ru/article/view/13608

Correspondence:
Marina Ivanova
ivanova_ma@pfur.ru

Received: December 29, 2021
Accepted: September 17, 2022
Published: September 30, 2022

INTRODUCTION

The cognitive system of an individual is most fundamentally determined by the level of its literacy (Bagiyan et al., 2021) both at the levels of oral and written speech. Writing skills include all the knowledge and abilities related to expressing ideas through the written words (words, spelling, grammar, punctuation). Correct spelling, grammar and punctuation are key factors in written communication. The readers will form an opinion of the author based on both the content and presentation. Errors in the written text are likely to lead them to form a negative impression. When researchers examine writing skills, as a rule, they consider and evaluate them as a complex phenomenon (Graham & Harris, 2011; Hardman & Bell, 2019), and few studies are devoted separately to the spelling rules and to the use of punctuation marks (Al-Tamimi, 2018; Ali et al., 2020; Tavşanlı & Kara, 2021).

Being an integral part of any written text punctuation not only reproduces the intonation of colloquial speech, but also serves as a mean of text structuring, making them understandable for the reader. Punctuation lets avoid ambiguity; it organizes both the whole discourse and its individual elements. It is enough

to remember the classical examples: A woman, without her man, is nothing and A woman: without her, man is nothing. Correct punctuation is indisputably very important. If the author of the text uses punctuation effectively it means that the reader will be able to make sense of the writing and will understand author's ideas. Punctuation symbols can also help create a particular effect to give depth to the text. In this regard, punctuation deserves to be examined separately, including academic writing.

Different issues on academic writing are raised in several studies: grammar and lexical bundles (Akbulut, 2020; Birhan, 2021), methodology for teaching academic writing (Kwak, 2017; Hussain, 2019), error analysis (Benjamin, 2017), digital resources support (Åberg et al., 2016; Alresheedi, 2019; Bezus et al, 2021; Strobla et al., 2019), etc. However, the researchers draw little attention to the punctuation as a particular side of academic writing skills (Al-Tamimi, 2018; Ali et al., 2020; Alharthi, 2021). As a rule, they explore the results of poor knowledge of punctuation, i.e. the different types of punctuation errors made by students in their writings. The methodology and digital support in punctuation training have not been investigated meticulously. The present study is an attempt to fill in this gap.

This paper is aimed to find the place of digital resources when teaching punctuation in academic writing course. The experimental comparative study was held to ascertain the possible influence of digital instruments on the students' achievements in punctuation while studying Academic English.

LITERATURE REVIEW

Academic writing skills are one of the key skills in higher education. According to K. Tusting and al., writing works are "at the heart of knowledge production" (Tusting, 2019). But mastering academic writing in English is the utmost complex and challenging task for *English as a Foreign Language* (EFL) students. Nevertheless, writing is the second most sought out skill after speaking in the EFL and academic English context (Alharthi, 2021).

As a rule, foreign learners face more problems with writing in English, which detract their essay content and also damage the harmony and esthetic of their writing (Williams, 2001). According to Williams, second language and foreign language learners encounter such critical language issues as vocabulary misuse, wrong use of suffixes and prefixes, grammatical errors, spelling, capitalization, punctuation and some others. The correct use of punctuation is very important in academic writing because punctuation marks bring clarity to the text. The problem is that the rules of punctuation are not taught to students properly and they keep on making errors in it.

Various studies have focused on how foreign students can improve their academic writing skills in English. Thus, R. Esfandiari & F. Barbary (2017), A. T. Birhan (2021) propose to develop students' academic writing skills with the help of lexical bundles.

The positive effect of online learning on writing skills was demonstrated in the work of P.J. González (2019). The researcher carried out an experimental comparative study in two groups of university students. One of the groups followed a face-to-face method, whereas another followed an online one. The findings of the experiment showed that the on-line group achieved better results in the written skills. The study of Baranovskaya et al. (2021) illustrated the fact that peer assessment processes greatly contribute to the development of students' 'academic writing skills'. Peer feedback resides in increasing ability to provide analyses of both, other students' essays and of their own ones. It was found out that collaborative peer assessment can help students to identify strengths and weaknesses in their own works and in the works of other classmates. The Schcolnik's research gives some recommendations on usage digital tools when writing in foreign language (Schcolnik, 2018). The study is focused on the digital tools that can support the professional process of academic writing (search engines, online libraries, phrase banks, dictionaries, grammar checkers, spellcheckers, etc.). The impact of the free-writing journal on foreign learners was examined in Alharthi's study (Alharthi, 2021). Five major areas were investigated to evaluate students' progress: the number of words written, spelling, capitalization, subject-verb agreement, and punctuation. The author has found that it is free-writing that improved students' writing skills.

Thus, the researchers examine different aspects of effective and correct academic writing in English. However, there are practically no works on methods and techniques of improving punctuation skills and abilities. Nevertheless, punctuation errors are the most dominant errors found in different writings of the foreign students (Sajid & Siddiqui, 2015). The analysis of the punctuation errors in Ali et al.'s research showed that the major reasons behind the differences in the frequency of punctuation errors in English were the lack of practice on punctuation marks, lack of teaching punctuation within context, lack of checking punctuation errors and lack of corrective feedback on the punctuation errors (Ali et al., 2020).

The present study seeks to answer the following questions to arrive at the above objectives: (1) What is the main punctuation issues which Russian students encounter in academic writing in English? (2) Can digital resources help reduce punctuation errors in academic writing? (3): What are the students' perceptions of digital support for learning punctuation in academic writing?

METHODS

Due to the modern time challenges and the demand for e-learning, traditional types of academic writing instruction in the classroom are not sufficient and need to be supported by digital resources. The present research summarizes the results of an empirical study: teaching academic writing punctuation in two groups of students who studied the same subject, *Academic English*, with the same teacher at the same time. One group (the control one) was applied a face-to-face and a text-book based methodology (the constant parameter). Another group (the experimental one) was trained not only with a face-to-face and a text-book methodology but also with digital support. The variable parameter of the study is a digital support applying method followed by one of the groups during the academic writing training process. The aim was to determine the presence or absence of significant differences in the results achieved.

Participants

We created the experimental group (EG), in which both the constant and variable parameters were presented, and the control group (CG), in which the variable parameter was absent. The EG consisted of 12 students; the CG consisted of 12 students as well. All the participants of the experiment were the third year students. The age of the students varied from 20 to 24 years. The third-year students were chosen for the study because their language level corresponds to the level B2 in accordance with the Common European Framework of Reference (CEFR). Thereby, they have sufficient experience in academic writing, in general, and their language experience would help them use digital resources in their independent additional practice. All the participants knew about the aims of the experiment and took part in the study voluntarily.

The training experiment was held on the basis of the Peoples' Friendship University of Russia during the second semester of the 2020-2021 academic year (Figure 1).

The Table 1 compares the main parameters of teaching methodology in the two groups.

Assessments and Measures

To get the most possibly accurate results of this study, different types of instruments have been chosen (Figure 1), namely: three final tests (measuring knowledge and abilities in punctuation; see Appendix C), a teacher's diary (Appendix D) and students' questionnaires (finding out their opinion and perception of the experiment; see Appendix A). The first instrument provides with quantitative information (RQ1), whilst the next two provide qualitative information (RQ2 and RQ3).

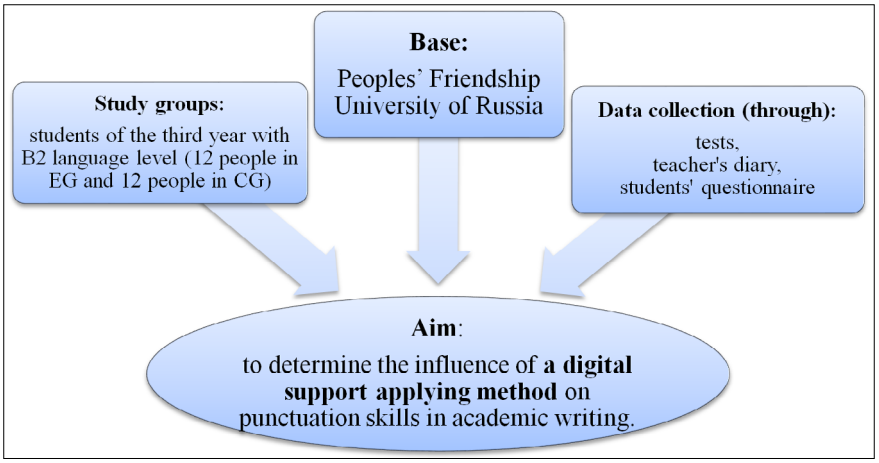
Table 1

Teaching methodology

CG	EG
Punctuation rules: teacher's explanations, examples from the Student's book.	Punctuation rules: teacher's recommendations on training videos, text digital resources, digital constructors, etc.
Making notes from the teacher's words	Making notes from different sources
Exercises from the Student's book and those created by the teacher	Different types of exercises (demonstrative, interactive, tests, quiz, etc.)
Evaluation and checking results during mutual discussion	Self-evaluation, checking here and now

Figure 1

Components of the training experiment



Teaching Instruments

The two groups were being trained academic writing punctuation for 10 weeks (one lesson of 2 academic hours' duration a week) with a face-to-face method using the Oxford textbook of the New Headway series *Academic Skills. Level 3* by Sarah Philpot and Lesley Curnick as required by university academic program. It should be noted that there are very few exercises on punctuation in this Student's book. They are mainly exemplary in nature, e.g. (p. 63):

When you quote directly from a source, you must use quotation marks. For example:

Carter (2006) believes, "The benefits to a city of holding the Olimpics will always outweigh the disadvantages" (p. 10).

When you quote indirectly, do not use quotation marks. For example:

Cox (2006) suggested that without adequate sponsorship, the Games could not take place.

and little training. e.g. (p. 63):

Rewrite the following direct quotations with the correct punctuation.

1. *according to Woodbridge 2004 p23 football is bigger business than selling food – According to Woodbridge (2004), "football is bigger business than selling food" (p. 23).*
2. *Khalil 2003 claims hosting large events usually leaves the host city with large bills p54*
3. *as Li Chung wrote 1999 the opportunity for development is considerable p71*

From our point of view such exercises are aimed to control the skills already achieved than to improve and to consolidate the abilities.

The themes to study during the experiment are: (1) The full stop and capitalized letters; (2) The comma; (3) The apostrophe; (4) Colon; (5) Semicolon; (6) Other punctuation symbols; (7) Quotation marks.

Due to the lack of exercises training punctuation skills in the Student's book, the teacher has to create additional exercises and explain punctuation rules (CG students made notes from teacher's words). Here are examples of some complementary exercises to support the explanation and a face-to-face teaching:

(1) Look at the following in-text citation, using an in-text citing system, in an essay.

There are also potential benefits to the use of artificial intelligence in medical context. Researchers report that robotic surgery can be used to assist in areas such as eye surgery, which requires a high level of detailed work in a small area (PLos ONE, 2013).

There is an error with the in-text citation. What is the problem?

Check one option.

a. The writer has not included a page number.
b. The in-text citation should appear at the beginning of the sentence.
c. The date of publication is not necessary.
d. The writer has given the publication name rather than the author's name. **(✓ correct)**

(2) Which pieces of information would you need to include with a direct quote from this source?

> Noda, Y., Ida, Y., Tanaka, S., Toyama, T., Roggia, M.F., Tamaki, Y., Sugita, N., Mitsuishi, M., & Ueta, T. (2013). Impact of robotic assistance on precision of vitreoretinal surgical procedures. PLoS ONE 8(1): 1-6. doi:10.1371/journal.pone.0054116.

Check three options.

a. PLoS ONE (journal name)
b. Noda, et al. (author surname) (✓ *Including the first author's surname (as well as any other authors if there are any, as demonstrated) is needed*.)
c. the page where the direct quote appears (✓ *For a direct quote, you always need to include the page number (or paragraph if on a website without pages).*)
d. robotic assistance (key words)
e. 2013 (year of publication) (✓ *For most referencing styles, such as APA 6th, you need to include the year of publication.*)

(3) When do we use square brackets in a direct quote?

Check one option.

a. To indicate that words or letters have been changed. **(✓ correct)**
b. To bracket the in-text reference.
c. To indicate that words have been removed.
d. To show which parts of the quote are less important.

To do the exercises the CG students were let time to revise the rules (from their notes), to look through the patterns in the Student's book, to find the examples in real academic texts, etc. They had no opportunities to check their answers immediately but only during the mutual discussion.

The EG training was supported by the main digital resource *Using English for Academic Purposes for Students in Higher Education. Academic Writing* by Andy Gillett (URL: http://www.uefap.com/writing/writfram.htm) containing different ICTs (Information and Communication Technologies) facilities in punctuation learning: interactive content, exercises, immediate evaluation, references, etc. The students were also

offered some additional training resources: *Academic Punctuation* (URL: https://academic-englishuk.com/punctuation/) and *Online Writing Lab* (URL: https://owl.purdue.edu/owl_exercises/punctuation_exercises/index.html) and some others (see the list in Appendix D).

The modern students are more accustomed to use digital devices than printed books therefore they prefer to receive information from digital resources. Taking it into account, the EG students were given the home tasks to revise the punctuation rules and to make notes from the short training videos (e.g., https://www.youtube.com/watch?v=aSVzzaamN7U&ab_channel=AcademicEnglishHelp) and from different text digital resources (e.g., https://academic-englishuk.com/punctuation/ and others).

Each digital resource chosen for the training contains variable punctuation exercises. E.g. (https://academic-englishuk.com/punctuation/):

- *Punctuate the following sentences:*

(1) vegware is an innovative company which produces compostable foodservice packaging from plants
(2) IBM is one of the worlds most ethical companies.

- *Error correct these sentences.*

- *Rewrite the paragraph with the correct punctuation*.

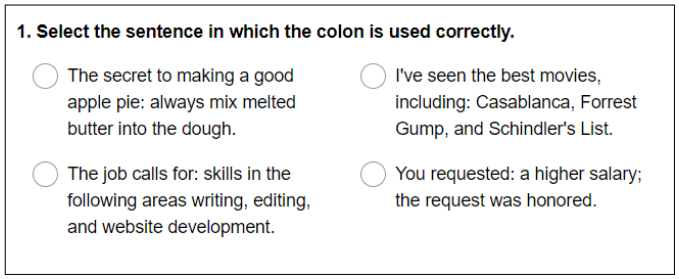
Pop-up responses provide an opportunity to save time and to get immediate results. The digital tests used in the lessons allow the teacher to save time for evaluating (Figure 2).

Self-Test Quiz lets the students not only to train different punctuation cases but also *here and now* to revise the appropriate punctuation rules by pressing the special button if the answer is incorrect (Figure 3).

The interactive window explains why the chosen answer is incorrect (providing with a ✓ *Rule to Remember*), gives the correct answer and offers the ways to correct the sentence.

Figure 2

Example of Advanced Punctuation Quiz



Resource: Advanced Punctuation Quiz. URL: https://www.proofreadnow.com/blog/advanced-punctuation-quiz

Such way of learning and revising the material involves various receptors and provides a complex effect on the students.

The *Citation styles* can be trained with the help of a "citation constructor" [1] (Figure 4).

To change the citation style it is enough to press the button "*Citation style*" and to choose other style, e.g. Chicago (Figure 5).

Thus, using this digital "citation constructor" the EG students could immediately see the changes in punctuation, revise the appropriate rules and make the appropriate notes.

Evaluation Tests

The present study comprises two sets of the experiment results: the results of the pre-test (pre-experiment stage) and the results of the final tests (post-experiment stage).

The purpose of the pre-experimental test was to reveal students' level in academic writing punctuation (Appendix B). The participants filled in a questionnaire in which they self-check their knowledge regarding the punctuation achieved before the experiment. This is what we have called the "pre-test".

The purpose of the final tests was to compare the level of improved punctuation skills in academic writing between the two groups. We have designed three tests. Each test consists of 5, 15 or 20 items respectively (Appendix C). The tasks were like these:

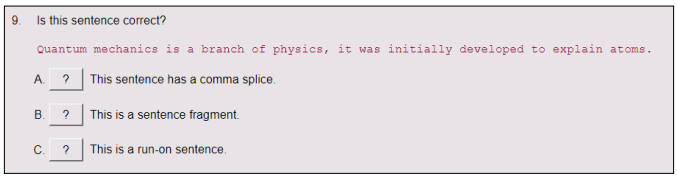
(1) Guess on what punctuation symbol is meant;

(2) Find errors in punctuation;

(3) Rewrite the text (or sentences) where the punctuation is removed completely.

Figure 3

Example of Punctuation Self-Test Quiz



Resource: Punctuation Self-Test. URL: https://www.niu.edu/writingtutorial/punctuation/quizzes/PunctuationSelfTest.htm

[1] Citation constructor. https://www.citethisforme.com

Figure 4

"Citation constructor": APA citation style

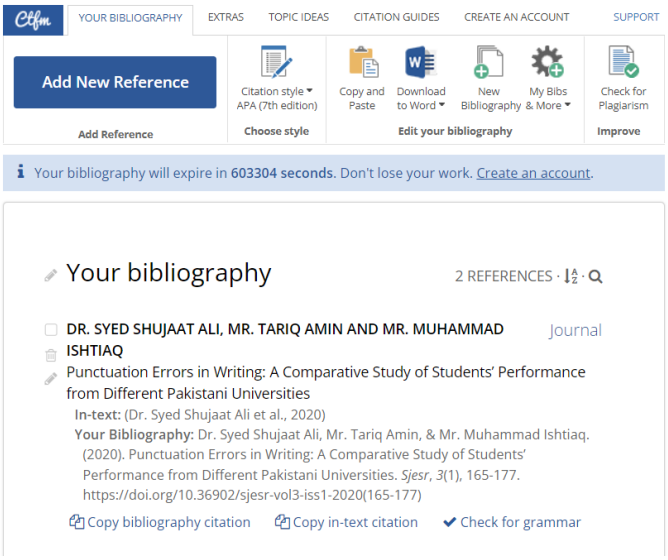
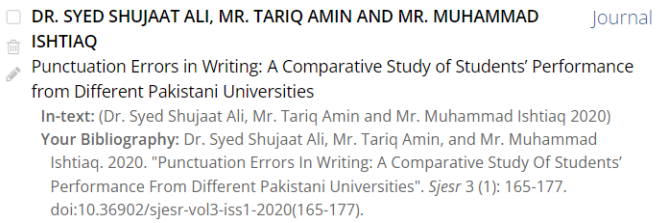


Figure 5

"Citation constructor": Chicago citation style



Teacher's Dairy

Another instrument that was used for this research is a teacher's diary in which we recorded the actions (themes, types of exercises, students' rating, etc.) carried out with each group (Appendix D). This information confirms that the lesson content of both groups was identical. The teacher's diary data would allow to know in details the difference between the two teaching methods, i.e. internal factors of the training process: dates, number of lessons, themes, home tasks, number of exercises done in the lesson, level of answers, etc. External factors are collected from the questionnaire.

Questionnaire

To study additional factors, that may influence the results of the experiment, a questionnaire has been designed in order to obtain students' opinions, perceptions and attitudes towards training process as well as their particular views on this experiment. Each item of the questionnaire is to be answered with a verbal Likert scale. We have also included several open-ended questions that could enrich the results. On the one hand, open-ended questions give students the opportunity to express their opinions more freely; on the other hand, such questions allow getting some qualitative information to complete the conclusions.

This is the final questionnaire (Appendix A) they are asked to fill in:

Please complete this questionnaire with your personal impressions of the series of lessons you have just assisted. The questionnaire is completely anonymous, so we ask you to be as honest as possible. The numbers correspond to: 1 – Strongly agree; 2 – Agree; 3 – Neither agree nor disagree; 4 – Disagree' 5 – Strongly disagree.

RESULTS

The analysis of the information provided by our students will yield conclusions about the use of ICTs or digital support in teaching academic writing punctuation.

Quantitative Data

The Table 2 provided us with the following data: a) the students know rules of punctuation marks using; b) the students are not sure about punctuation marks using; c) the students do not know punctuation marks using.

We can see that the simplest punctuation symbols for using are *a full stop, a comma* and *an apostrophe*: 80-100% of students are sure they know the rules.

The main punctuation issues that Russian EFL students encounter when writing academic English (RQ1) are *a colon, a semi-colon, a dash, a hyphen* and *brackets*: only 33-50% of the students have a clear understanding of the use of these punctuation marks. Thus, these punctuation signs are to be paid close attention in academic writing classes.

The Table 3 shows the final tests results: the figure in the box corresponds to the number of correct answers in the test.

According to the evaluation scale, developed by the Soviet methodologist V. Tsetlin for spelling skills, the level of punctuation skills can be also assessed as 'excellent' if 90–100% of the material is learned; as 'good' if 75–< 90% of the material is learned; as 'satisfactory' if 60–< 75% of the material is learned; as 'unsatisfactory' if less than 60% of the material is learned.

The first test (theoretical one), consisting of five items, did not cause any particular difficulties in both groups. All the students achieved 100% result. The second test (Table 4), consisting of 15 items, showed slight variation in results between CG and EG.

The third test (Table 5), consisting of 20 items, the most difficult one, was done successfully by most students from EG:

Table 2

Pre-Test Results

	CG (12 people)			EG (12 people)		
	a) I know when use	**b) It rings a bell of using**	**c) Never heard of using**	**a) I know when use**	**b) It rings a bell of using**	**c) Never heard of using**
a full stop	12	0	0	12	0	0
a comma	11	1	0	10	2	0
a colon	4	5	3	5	5	2
a semi-colon	4	5	3	5	4	3
an apostrophe	10	2	0	11	1	0
a dash	5	5	2	6	4	3
a hyphen	5	5	2	5	5	2
brackets	4	7	1	5	6	1

Table 3

Post-Test Results

CG Students	Test 1 (5 items)	Test 2 (15 items)	Test 3 (20 items)	Test 1 (5 items)	Test 2 (15 items)	Test 3 (20 items)	EG Students
1	5	15	18	5	15	20	1
2	5	14	18	5	15	20	2
3	5	13	17	5	14	19	3
4	5	13	16	5	13	18	4
5	5	12	16	5	13	17	5
6	5	12	15	5	12	17	6
7	5	11	15	5	12	17	7
8	5	11	15	5	11	17	8
9	5	10	14	5	10	16	9
10	5	10	13	5	10	15	10
11	5	9	12	5	10	15	11
12	5	9	10	5	9	14	12

Table 4

Second Test Results

	excellent	**good**	**satisfactory**	**unsatisfactory**
Control group	16,7%	50%	33,3%	0
Experimental group	25%	41,7%	33,3%	0

Table 5

Third Test Results

	excellent	**good**	**satisfactory**	**unsatisfactory**
Control group	16,7%	50%	25%	8,3%
Experimental group	33,3%	58,4%	8,3%	0

Thus, we can join the opinion of the researchers who believe that digital support can improve academic writing skills, in particular it can help reduce punctuation errors in academic writing (RQ2).

Qualitative Data

The qualitative data of this experiment were drawn from the *Personal Questionnaire* that our students have filled in. The table given in the Appendix A presents the EG students' opinions, perceptions and attitudes towards training process according to a verbal Likert scale (RQ3).

The summary and conclusions of the data from the *Personal Questionnaire*:

- Most of the student, by their opinion, had some previous knowledge about academic writing punctuation. Three punctuation marks (*a full stop, a comma* and *an apostrophe*) are the simplest and are known by 80-100% of students. Five punctuation symbols (*a colon, a semi-colon, a dash, a hyphen* and *brackets*) require closer attention when teaching academic writing for Russian EFL students because only half of the experiment participants have clear ideas of using these punctuation signs.
- Most of the students consider that the text-book is not enough to learn punctuation; the study is to be supported by digital resources.
- All of them stated that they have improved their knowledge about academic writing punctuation and they will be able to use punctuation symbols correctly.

The Table 6 presents the most common, given in a free manner answers of the EG students about the training exercises (RQ3).

The qualitative data tell us that the great majority of the students consider the experiment as positive and useful. In spite of the fact that they had their difficulties and their dislikes we can say that the experience should be repeated with other academic groups (e.g., of different levels) to get more experimental material.

DISCUSSION

The present study has contributed to a better understanding of the development of upperintermediate undergraduates' written competence. Written competence is a subset of learners' language competence with an emphasis on writing-specific abilities including punctuation. The correct punctuation is very important, as for academic writing it is important twice. Nevertheless, there are scarce researches on the problem of using punctuation. Meanwhile punctuation errors occupy the 4th place (after diction, poor expression and prepositions) in Academic Writing in English Language (e.g. among Pakistan students) according to M. Sajid & A. Siddiqui's analysis.

In their turn S. Ali et al. claim that students are not taught punctuation marks in English as a foreign language that leads to numerous errors. Our findings do tally with their opinion that the absence of special lessons or courses is one of the basis reasons behind punctuation errors. Our study shows one of the possible ways of filling this gap by organizing a series of specialized lessons on punctuation in academic writing course with digital resources. The *Questionnaire* has confirmed that such an approach raises interest to the training. While the research of Tavşanlı & Kara demonstrates another method of creating a positive attitude toward writing, increasing motivation, learning desire and communication skills: it is peer and self-assessment that have a positive effect on the correct use of spelling rules and punctuation marks.

Moreover S. Ali et al.'s study reveals different types of punctuation errors. The most frequent ones are capital, comma and apostrophe. "*Comma error is on top*" [Ali et al., 2020, 170]. Having done the content analyses, V. Halitoğlu also notes that the most common punctuation errors among Turkish students are errors in the use of commas; the less frequent ones are the use of apostrophe. S. Ali et al.'s consider that the comma errors are found to be the most frequent be-

Table 6

Open-Ended Questions

What exercises were the most helpful for you?	Find punctuation errors in the sentences and correct them. / Advanced punctuation quiz.
What exercises were the least helpful for you?	None. / Every exercise is aimed to train different aspects.
What exercises were the most difficult for you?	Restore missing punctuation. / Restore capital letters in the text. / Citation. / Quotation.
What exercises were too easy for you?	Select the sentences in which … is used correctly/incorrectly. / Multiple choice tests.
Other comments	I think that this way of learning punctuation is very useful. / I liked this course because it has put in order my knowledge on academic writing punctuation. / I have known several new rules on punctuation use.

cause of its various functions in texts. It is said that there are fourteen functions of comma. The researchers recommend arranging different seminars on the importance of punctuation marks.

The present research doesn't focus on types but compares the total number of punctuation errors in the Russian students' test works of two groups. Although the Russian students themselves note that they do not quite know the rules of using colons, semicolons, dashes, hyphens and brackets. Our experiment reveals that the digital support can help diminish the number of punctuation errors. Meanwhile Thewissen presents punctuation errors as improvement resistant features across proficiency levels. A.C. Lahuerta doesn't agree with that and claims that punctuation errors seem to show strong association with certain proficiency levels. Her study reveals that these error categories reduce as proficiency increases.

In further researches, it is necessary to identify not only the relationship between digital support and the total number of punctuation errors in academic written production, but also error types as well as confusion between punctuation markers, missing or redundant markers.

CONCLUSION

Whether we like it or not, digit and digital learning have forever penetrated in our everyday life and in education system, in particular. COVID-19 has made significant changes in the teaching process and there are already few attempts to raise a question of the impregnability of such a form of interaction as *teacher* <–> *student*. In this, in the recent past, two-link chain, one more link is being integrated – digital educational resources: *teacher* <–> *digital educational resources* <–> *student*. The effectiveness of using digital resources in teaching various aspects of academic writing has been proven in the works of many researchers. In the present one, the authors have made an attempt to reveal the features of working with digital resources when teaching punctuation in academic writing as one of the sides of writing skills.

ICTs are not the panacea for the learning process. But the experiment and our pedagogical experience have demonstrated that digital resources are a good instrument to be used by the language teacher. Many students evaluated the experiment as a positive one. They enjoyed working with digital resources, they need interactivity and they stressed that a usual printed Students' book is not enough to form firm punctuation skills.

Thus, digital support should be one of the main parts of the academic writing training process and should be included in training programs. The materials of the present paper (resources on punctuation, evaluating tables, types of tasks and tests) can be used in a real teaching practice. We suppose that digital support can be successfully applied in university practice with the elementary and advanced level students. In further researches it is necessary to take into account not only the number of punctuation errors but also their types and frequency of use.

DECLARATION OF COMPETING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Marina Ivanova: conceptualization, data curation, investigation, methodology, project administration, validation, visualization, writing-original draft, formal analysis.

Nadezhda Arupova: conceptualization, formal analysis, methodology, project administration, supervision, validation, visualization, writing-original draft, writing-review & editing.

Natalya Mekeko: conceptualization, methodology, project administration, resources, supervision, validation, writing-review & editing.

REFERENCES

Åberg, E. S., Ståhle, Y., Engdahl, I., & Knutes-Nyqvist, H. (2016). Designing a website to support students' academic writing process. *Turkish Online Journal of Educational Technology - TOJET*, *15*(1), 33–42.

Alharthi, S. (2021). From Instructed writing to free-writing: A study of EFL learners. *SAGE Open*, 1–15. https://doi.org/10.1177/21582440211007112

Ali, S. S., Amin, T., Ishtiaq, M. (2020). Punctuation errors in writing: A comparative study of students' performance from different Pakistani Universities. *Sir Syed Journal of Education & Social Research*, *3*(1), 165-177. https://doi.org/10.36902/sjesr-vol3-iss1-2020(165-177)

Alresheedi, H. (2019). *Investigating the efficacy of computer and mobile assisted vocabulary learning in developing academic writing skills with Saudi university students* [Doctoral dissertation]. TEFL 5000-Dissertation. https://files.eric.ed.gov/fulltext/ED610156.pdf

Al-Tamimi, N. (2018). An investigation of academic writing errors among tertiary level students at Hadhramout university: A perception study. *Journal of Applied Linguistics and Language Research*, *5*(2), 214-227.

Akbulut, F. D. (2020). A bibliometric analysis of lexical bundles usage in native and non-native academic writing. *Journal of Language and Linguistic Studies*, *16*(3), 1146–1166.

Bagiyan, A., Shiryaeva, T., Tikhonova, E., & Mekeko, N. (2021). The real value of words: how target language linguistic modelling of foreign language teaching content shapes students' professional identity. *Heliyon, 7*(3), e06581. https://doi.org/10.1016/j.heliyon.2021.e06581

Baranovskaya, T. A., Khorokhorina, G. A., & Son, L. P. (2021). The effect of peer assessment on the quality of written performance in Distance English Language Learning (by the example of academic writing training course in Russian university). In *Proceedings of the 12th International Conference on E-Education, E-Business, E-Management, and E-Learning* (pp. 105–112). https://doi.org/10.1145/3450148.3450205

Benjamin, A. (2017). Error analysis of students essays: A case of first year students of the university of health and allied sciences. *International Journal of Higher Education*, *6*(4), 54–68.

Bezus, S., Mekeko, N., & Akbilek, E. (2021). Digital Resources for Learning and Teaching Academic Writing in English. *12th International Conference on E-Education, E-Business, E-Management, and E-Learning* (pp. 53–59). IEEE. https://doi.org/10.1145/3450148.3450206

Birhan, A. T. (2021). Effects of teaching lexical bundles on EFL students' abstract genre academic writing skills improvement: Corpus-based research design. *International Journal of Language Education*, *5*(1), 585–597. https://doi.org/10.26858/ijole.v5i1.14917

Esfandiari, R., & Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purpose*, 29, 21–42. https://doi.org/10.1016/J.JEAP.2017.09.002

González, P. J. (2019). Estudio comparativo experimental entre un entorno virtual y otro presencial en una asignatura de Inglés [Experimental comparative study between a virtual environment and a face-to-face environnent in an English subject]. *Porta Linguarum*, *32*, 119–134.

Graham, S., Harris, K., & Hebert, M. A. (2011). *Informing writing: The benefits of formative assessment. A Carnegie corporation time to act report*. Alliance for Excellent Education. https://media.carnegie.org/filer_public/37/b8/37b87202-7138-4ff9-90c0-cd6c6f2335bf/ccny_report_2011_informing.pdf

Halitoglu, V. (2021). Determination of the Writing Errors of Bilingual Turkish Students in France (Example of Lyon). *Participatory Educational Research*, *8*(4), 297–320. http://dx.doi.org/10.17275/per.21.91.8.4

Hardman, W., & Bell, H. (2019). More fronted adverbials than ever before'. Writing feedback practices and grammatical metalanguage in an English primary school. *Language and Education*, *33*(1), 35–50. https://doi.org/10.1080/09500782.2018.1488864

Hussain, S. S. (2019). Strategies for teaching academic writing to Saudi L2 learners. *English Language Teaching*, *12*(12), 1–11. https://doi.org/10.5539/elt.v12n12p1

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*(2), 112–133. https://doi.org/10.1177/1558689806298224

Kwak, S. (2017). Approaches reflected in academic writing MOOCs. *The International Review of Research in Open and Distributed Learning*, *18*(3). https://doi.org/10.19173/irrodl.v18i3.2845

Lahuerta, A. C. (2018) Study of Accuracy and Grammatical Complexity in EFL Writing. *International Journal of English Studies*, *18*(1), 71–89. doi:10.6018/ijes/2018/1/258971

Sajid, M. & Siddiqui, A. (2015). Lack of academic writing skills in English language at higher education level in Pakistan: Causes, effects and remedies. *International Journal of Language and Linguistics*, *2*(4), 174–186.

Schcolnik, M. (2018). Digital tools in academic writing. *Journal of Academic Writing*, *8*(1), 121–130. https://doi.org/10.18552/joaw.v8i1.360

Strobla, C., Ailhaudb, E., Benetosc, K., Devittd, A., Krusee, O., Proskef, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education*, 131, 33–48. https://doi.org/10.1145/3450148.3450206

Tavşanlı, Ö. F., & Kara Ü. E. (2021). The effect of a peer and self-assessment-based editorial study on students' ability to follow spelling rules and use punctuation marks correctly. *Participatory Educational Research* (PER), *8*(3), 268–284. http://dx.doi.org/10.17275/per.21.65.8.3

Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: Insights from an error□tagged EFL learner corpus. *The Modern Language Journal*, 97, 77–101.

Tusting, K., McCulloch, S., Bhatt, I., Hamilton, M., & Barton, D. (2019). Academic writing. The dynamics of knowledge creation. London: Routledge. https://doi.org/10.4324/9780429197994

Williams, J. A. (2001). Classroom conversations: Opportunities to learn for ESL students in mainstream classrooms. *Embracing Pluralism Worldwide, 54(8), 750–757. https://www.jstor.org/stable/20204989*

APPENDIX A

	1	2	3	4	5
I have already had prior knowledge of the academic writing punctuation.	9	3	0	0	0
It is possible to learn punctuation only with the help of the student's book in traditional classroom.	0	1	2	9	0
It is better to learn punctuation with digital support.	12	0	0	0	0
The participation of a tutor/teacher is necessary.	12	0	0	0	0
The course has helped me improve my knowledge in academic writing punctuation.	10	2	0	0	0
I would be able to use the punctuation symbols learned in real contexts.	11	1	0	0	0
I have gotten high level in training exercises.	10	1	1	0	0
Using the interactive programs is quite useful.	9	3	0	0	0
The time of training course is sufficient.	7	2	3	0	0
I positively evaluate the training experience.	10	2	0	0	0

APPENDIX B

pre-experimental test

	a) I know when use	b) It rings a bell of using	c) Never heard of using
a full stop			
a comma			
a colon			
a semi-colon			
an apostrophe			
a dash			
a hyphen			
brackets			

APPENDIX C

Final tests

Test 1. Which punctuation symbol is meant? Write the appropriate name in the gaps.

1. Use a … to create a pause, to separate ideas in the sentence.
2. Use a … to create a break, but recognize connection of ideas.
3. Use a … to connect two sentences thematically.
4. Use a … to create the end of the sentence.
5. Use an … to indicate ownership or missing letters/numbers.

Answers: *1) comma; 2) semi-colon; 3) colon; 4) full stop; 5) apostrophe.*

Test 2. Find missing or incorrect punctuation.

1. Experience indicates that, these problems do not result wholly from our lifestyle.
2. The second exercise which was more difficult took half an hour.
3. A man of his great abilities, would always be successful.
4. The number of service enterprises in wealthier free-market economies, has grown rapidly.
5. The policy on the single European currency is not well defined, it confuses many people.
6. Terrorism legislation will be discussed alongside the Human Rights Act, this is because they contradict each other, this is a crucial matter.
7. Recently the number of service enterprises in wealthier free-market economies has grown rapidly.
8. Some businesses only seek to earn enough to cover their operating costs however.
9. A first-aid kit should contain the following items; cotton wool, sticking plasters, antiseptic cream, bandages and a pair of scissors.
10. The sciences search for change, the humanities reflect on what it is.
11. An individuals right to security is guaranteed.
12. Security is very important especially for large corporation's.
13. Each application has it's own advantages and disadvantages.
14. Leonardo da Vinci, 1452–1519, was an advocate of vegetarianism.
15. Introduction:

Answers:

1. Experience indicates that these problems do not result wholly from our lifestyle. (NB: Commas cannot be used after *that*.)
2. The second exercise, which was more difficult, took half an hour. (Two commas are needed to set off the extra information.)
3. A man of his great abilities would always be successful. (To separate the subject (who or what) from the verb is a mistake.)
4. The number of service enterprises in wealthier free-market economies has grown rapidly. (See № 3)
5. The policy on the single European currency is not well defined. It confuses many people.
6. Terrorism legislation will be discussed alongside the Human Rights Act. This is because they contradict each other. This is a crucial matter.
7. Recently, the number of service enterprises in wealthier free-market economies has grown rapidly.
8. Some businesses only seek to earn enough to cover their operating costs, however.
9. A first-aid kit should contain the following items: cotton wool, sticking plasters, antiseptic cream, bandages, and a pair of scissors.
10. The sciences search for change; the humanities reflect on what it is.
11. An individual's right to security is guaranteed.
12. Security is very important, especially for large corporations.
13. Each application has its own advantages and disadvantages
14. Leonardo da Vinci (1452–1519) was an advocate of vegetarianism.
15. Introduction

Test 3. Try reading this. Then restore the punctuation and capitalized letters.

while the country saw a certain amount of religious freedom during the first half of the XVIIth century the latter was notable for several attempts to homogenize religious practice through legislation which was a response to the fears of an anglican majority who anticipated subversion of interest to the history of literary studies are the penal laws that made holy communion according to the rites of the church of england a condition precedent for entrance to the sole universities in the country at the time oxford and cambridge embodied in 1661 as laws in the form of the religious tests acts they directly led to the establishment of the first higher education institutions (or *academies*) in the country to teach the concerted study of english literature by the communities who refused to conform to the anglican church and who were therefore precluded from higher education known as the *Dissenters*.

Answer:

(1) While the country saw a certain amount of religious freedom during the first half of the XVIIth century (2), the latter was notable for several attempts to homogenize religious practice through legislation which was a response to the fears of an (3) Anglican majority who anticipated subversion (4). (5) Of interest to the history of literary studies are the penal laws that made holy communion according to the rites of the (6) Church of (7) England (8), a condition precedent for entrance to the sole universities in the country at the time (9): (10) Oxford and (11) Cambridge (12). (13) Embodied in 1661 as laws in the form of the (14) Religious (15) Tests (16) Acts (17), they directly led to the establishment of the first higher education institutions (or *academies*) in the country to teach the concerted study of (18) English literature by the communities who refused to conform to the (19) Anglican (20) Church and who were therefore precluded from higher education known as the *Dissenters*.

APPENDIX D

teacher's diary

Lesson number	1	2	3	4	5	6	7	8	9	10
Student 1										
Student 2										
Student 3										
Student 4										
Student 5										
Student 6										
Student 7										
Student 8										
Student 9										
Student 10										
Student 11										
Student 12										

Lesson №	Theme	Digital support
1	The full stop and capitalized letters	1) http://www.uefap.com/writing/exercise/punc/puncex4.htm 2) https://academic-englishuk.com/punctuation/ 3) http://www.uefap.com/writing/exercise/punc/puncex8.htm
2	The comma	1) https://owl.purdue.edu/owl_exercises/punctuation_exercises/commas/index.html 2) http://www.butte.edu/departments/cas/tipsheets/practice_punctuation.html 3) https://www.youtube.com/watch?v=aSVzzaamN7U&ab_channel=AcademicEnglishHelp
3	The apostrophe	1) https://owl.purdue.edu/owl_exercises/punctuation_exercises/apostrophes/apostrophes_exercise.html 2) http://www.butte.edu/departments/cas/tipsheets/practice_punctuation.html
4	Colon	http://www.butte.edu/departments/cas/tipsheets/practice_punctuation.html
5	Semicolon	http://www.butte.edu/departments/cas/tipsheets/practice_punctuation.html
6	Other punctuation symbols	http://www.butte.edu/departments/cas/tipsheets/practice_punctuation.html
7	Quotation marks	1) https://owl.purdue.edu/owl_exercises/punctuation_exercises/quotation_marks/quotation_marks_exercise.html 2) https://academic-englishuk.com/punctuation/ 3) https://www.citethisforme.com
8	Mixed training 1	1) https://owl.purdue.edu/owl_exercises/punctuation_exercises/basic_punctuation/punctuation_exercise.html 2) https://www.bbc.co.uk/bitesize/guides/zw8vwxs/test
9	Mixed training 2	1) https://www.proofreadnow.com/blog/advanced-punctuation-quiz 2) https://www.niu.edu/writingtutorial/punctuation/quizzes/PunctuationSelfTest.htm
10	Final test	

https://doi.org/10.17323/jle.2022.12654

The Relationship between EFL Learners' Flipped Learning Readiness and their Learning Engagement, Critical thinking, and Autonomy: A Structural Equation Modelling Approach

Afsaneh Ravandpour

University of Tehran, Kish International Branch, Kish, Iran

ABSTRACT

Purpose. The main purpose of this study is to investigate the relationship between Iranian EFL learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy.

Methods. The participants comprised 520 EFL learners studying at various language institutes in the context of Iran. They were chosen according to convenience sampling. In this study, four instruments were used: flipped learning readiness questionnaire, learner autonomy scale, critical thinking inventory, and learning engagement questionnaire. To analyze the relationships among the variables, Pearson's correlation coefficient and structural equation modeling (SEM) was run. Results. The results revealed that flipped learning readiness correlated positively and significantly with three variables: learning engagement, critical thinking, and autonomy. In addition, based on the results, flipped learning is a positive significant predictor of critical thinking, learning engagement, and autonomy. Moreover, engagement is positively predicted by both critical thinking and autonomy. Finally, the results and their implications in the context of language learning were discussed.

Implications.

KEYWORDS

flipped learning readiness, learning engagement, learner critical thinking, learner autonomy, English as a foreign language (EFL) learner

INTRODUCTION

In the last decades, the development of technology has affected various fields of sciences including education dramatically. The goal of applying educational technologies to pedagogical settings is to improve the quality of teaching, instructional materials, and teaching methods to their utmost utility (Berrett, 2012). In particular, the integration of face-to-face classrooms with technological tools is a key factor in the enhancement of teaching and learning context in the 21st century (Arum & Roska, 2011; Graham et al., 2013; McLaughlin et al., 2014). Easy access to educational technologies has changed the way learners learn and the role instructors play (Johnson, Adams Becker, Estrada, & Freeman, 2015). Learners can now look for information online and instructors are no longer the single supplier of information and knowledge. In this new context, instructors have been looking for an instruction model which best meets their learners' needs. Among the most popular new educational technologies, flipped instruction has been considered as one such model (Bergmann & Sams, 2012).

Flipped instruction has been popularized as a creative learning practice for supporting teaching, especially language teaching. As a new form of blended learning, the flipped instruction reverses traditional instruction and rearranges the time of instruction to provide more

Citation: Ravandpour A. (2022). The Relationship between EFL Learners' Flipped Learning Readiness and their Learning Engagement, Critical thinking, and Autonomy: A Structural Equation Modelling Approach. Journal of Language and Education, 8(3), 97-105. https://jle.hse.ru/article/view/12654

Correspondence:
Afsaneh Ravandpour,
afsanehravandpour2018@gmail.com

Received: June 26, 2021
Accepted: July 10, 2022
Published: September 30, 2022

class time for learners to learn (Öztürk & Çakıroğlu, 2021). In this kind of instruction, assignments are assigned to learners before the classroom and the time of class is devoted to higher-order activities. Thus, in the flipped classroom, the content of instruction is introduced to the students using the technology outside the classroom context (Touchton, 2015 & Strayer, 2012). In this type of classroom, students would use asynchronous online lectures along with face-to-face interactions within the classroom (Bergmann & Sams 2013). Thus, the time of the class is devoted efficiently to scaffolding, inspiration, and assistance about the materials presented previously online to the students. Flipped instruction also provides language learners with more chances to be exposed to the L2 language both inside and outside the classroom by the technology-integrated mechanism of learner previews and in-class differentiated education (Bergmann & Sams, 2012).

There are many advantages about the flipped classrooms such as suggesting the individualized learning style for students (Green, 2015), presenting teachers with the customized curriculum (Herreid & Schiller, 2013), and covering various types of material that are not possible to be covered in traditional models (Mason, Shuman, & Cook, 2013). The inclusion of the students and engaging them in the process of learning and developing their critical thinking ability as well as problem-solving skills (Engin & Donanci, 2016; Green, 2015) are other benefits. Critical thinking, creative learning, and complex reasoning skills are the overlooked elements of education (Arum & Roska 2011; Graham et al. 2013; McLaughlin et al. 2014). Some researchers claimed that the application of new digital pedagogies has established a proper base for the development of these skills amongst students. Therefore, the flipped classroom is considered a way for the expansion of the curriculum rather than the instruction of the content in a digital format (Bishop & Verleger, 2013). A great benefit of flipped learning is that students could be engaged with some of the most important parts of the instruction content in advance. In order words, it could be said that teachers can give some hints about what they want to teach before the instruction before the learners attend the class so they can think about the material carefully (O'Flaherty & Phillips, 2015). Thus, studies on this issue are so significant for learners and teachers.

In flipped instruction, learners should take responsibility for their learning by implementing several assignments before the class. According to Han (2015), as the flipped classroom structure needs learners to be actively engaged in learning at the same time with leaner education, the enhancement of learner engagement as well as autonomy could be noticed. Moreover, flipped instruction empowers teachers to provide learners with chances to communicate with each other, as well as to motivate learners by holding discussions, answering questions, and modeling problem-solving activities (Bergmann & Sams, 2012). Thorndike (1932) argued that one rule of learning is the readiness of learners to learn and this can strongly affect the degree of learning achievement. Therefore, this study examined flipped learning readiness and its relation to learning engagement, critical thinking, and autonomy. Regarding the importance of flipped instruction in the EFL context, more studies should be devoted to this issue. Therefore, because of the significance and scarcity of studies in this field, this investigation aimed to engage in this significant issue. Thus, the main purpose of this study is to examine the relationship between Iranian EFL learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy.

Numerous efforts have been taken to examine the effect of technology in developing and designing instructional materials. These steps have been taken to change the traditional classes into pedagogical settings with a high-quality atmosphere. Previously, instructors carried the main burden of lecturing and assigning the homework for the students, while with the growth of online instruction tools (e.g., in the flipped approach) these roles have been changed, the students engaged more in their learning process, and new requirements came into existence. For instance, the issues traditionally discussed in the class are taking place outside the classrooms and online (Lage, Platt, & Treglia, 2000). Among different possibilities for implementing the flipped classroom, student-centered teaching model, promoting cooperative and active learning is one of the bests (Lee & Wange, 2013). The underlying theory for flipped classrooms is the constructivist principles, which asserts students' engagement cultivates their critical thinking activities. In Our changing and challenging world, learners need to go beyond the creation of their knowledge; they require to develop their higher-order thinking skills, including critical thinking and problem-solving (Sezer, 2008).

Among the studies that have been done regarding technology, Yarbro, Arfstrom, McKnight, and McKnight (2014) examined research on flipped learning from K-12 to post-secondary education and found that it can be utilized in most fields, such as math and English as a foreign language (EFL). They also concluded that learners in flipped classrooms were more engaged, act better, or showed fewer behavioral problems. In the case of flipped education and learners' autonomy, Han (2015) examined the effectiveness of flipped classrooms on ESL learners` autonomy based on Strayer's (2007) class-flipping model. According to his findings, flipped classrooms had a significant effect on ESL learners` autonomy. As flipped learning needs students to be actively engaged in learning, the development of learner autonomy could be achieved. Hao (2016) who considered the flipped learning readiness variable, investigated the impact of personal characteristics on their flipped learning readiness levels. According to the results, personal characteristics and individual circumstances affect the levels of readiness. As literature reviewed, the role of flipped learning readiness on learners' learning engagement, critical thinking, and autonomy simultaneously has not been examined yet. Thus, this

study aims to fill this gap and presents implications in the EFL context of language learning.

Considering the role of flipped instruction and technology in education, Moranski and Kim (2016) in another study, examined the effect of presenting grammar content by flipped instruction and technology and traditional methods on third-semester Spanish learners' (*N* = 213) knowledge of the non-agentive clitic pronunciation. They found that FI helps students automatize explicit knowledge, and on the other, that applying educational videos and readings does not negatively affect learners' proficiency to correctly explain complex target structures. Regarding flipped education and learners' engagement, Gasmi (2017) examined the impact of flipped instruction on EFL learners' engagement in academic writing classes in Oman and found this kind of instruction effective. His study represented new practicable knowledge regarding the implementation and implications of flipped learning for EFL learners' engagement in Oman in academic writing. This study presents flipped learning as an educational approach that helps to address learners' lack of behavioral, cognitive, emotional, and agentic engagement in writing courses in the context of Oman.

Considering the learners' attitudes Muniandy (2018, cited in Shahani, Chalak & Heidari Tabrizi, 2021) examined the effectiveness of flipped classrooms on learners` achievement and attitudes towards the English language in secondary education in Malaysia. His results indicated that flipped classroom affects positively learners' achievement and attitudes more than traditional classrooms. Also, the findings revealed that flipped classroom provides greater opportunities for communication between learners and their instructors. Moreover, Vaezi, Afghari, and Lotfi (2019) in their study by a mixed-method research approach examined EFL learners' and teachers' attitudes towards flipped classrooms in an Iranian university context. According to their findings, the learners had positive attitudes towards flipped instruction. Tecedor and Perez (2019) also investigated how learners interpret and experience their roles as students in flipped Spanish courses and how these conceptualizations shape their conduct and beliefs about learning under this new educational approach. They concluded that the major "predictors of liking FIs are (i) being enrolled in an elementary course, (ii) having taken flipped courses in other subject areas, (iii) having a clear understanding of how the online platform works, and (iv) understanding the philosophy behind flipped courses" (p.1). Finally, considering the impact of flipped education on learners' critical thinking, Viriyavejakul (2020) in a study investigated the impact of flipped instruction on undergraduate students' critical thinking and found a positive relationship between them.

The following research question was posed and was examined in this study: Are there any statistically significant relationships between EFL learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy?

METHODS

Participants

A total number of 520 EFL learners (300 female, 220 male) participated in this study from four private language institutes on Kish Island, Iran. Their selection was based on convenience sampling and the participation was entirely voluntary. Because of the lack of access to all Iranian EFL learners, random sampling was not possible and so, the participants were selected by convenience sampling. They were between the age range of 14 and 39 (mean=24.35, SD=4.24). All of them were intermediate learners of English and were able to fill the English version of the scales. Because of the Covid-19 situation, the learners took part in online classes and they took advantage of flipped classroom instruction, so they were familiar with this type of instruction and learning.

Instruments

In this study, one questionnaire, comprised of the following scales, was used: flipped learning readiness questionnaire, learner autonomy scale, critical thinking inventory, and learning engagement questionnaire.

Flipped Learning Readiness Questionnaire

Flipped learning readiness questionnaire was developed and validated by Hao (2016) and includes 27 items based on a 5-point Likert scale varying from "completely disagree" (1) to "completely agree" (5). The scale included 4 subscales on learner control and self-directed learning, technology self-efficacy, motivation for learning, in-class communication self-efficacy, and doing previews. "The face validity of these items was confirmed by two experts, who compared them to items in Smith's (2005) study" (Hao, 2016, p. 297). The reliability was reported with a sufficient alpha level ranging from .75 to .92

Learner Autonomy Scale

The learner autonomy Scale was devised by Sakai et al. (2008) and it was adopted as such in this study. This scale is used for examining the learners' autonomy. The questionnaire includes 65 items and assesses five dimensions of "Recognition of Responsibility for Learning, Responsibility for Past Learning, Responsibility for Future Learning, Past Learning outside Classroom, and Future Learning outside Classroom.". Responses were given a 5-point Likert scale ranging from (1) 1-Not at all, 2-Hardly, 3-To some extent, 4-Mostly, (5) Totally". This scale has high reliability and validity. The Cronbach's coefficient alphas were bigger than .80.

Critical Thinking Inventory

To measure learners' critical thinking California Critical Thinking Skills Test (CCTST) was used in this study. This test includes 34 multiple-choice questions developed to measure critical thinking using two sub-scales. The primary subscale consists of Analysis, Evaluation, Inference, and the secondary subscale includes Deduction and Induction (Sadeghi, Hassani, & Rahmatkhah, 2014).

Learning Engagement Questionnaire

This questionnaire was developed and validated by Gasmi (2017) which was a modification of three validated instruments (Greene's (2015) Cognitive Engagement Scale, Miserandino's (1996) Perceived Behavioral and Emotional Engagement Questionnaire, & Reeve and Tseng's (2011) students' agentic engagement). This questionnaire has 4 components measuring students' Agentic Engagement, Behavioral Engagement, Cognitive Engagement, and Emotional Engagement. This scale has 67 items. "Each item in the SEQ was measured on a 6-point Likert scale ranging from 'strongly agree' (6) to 'strongly disagree' (1) in Parts I, II, III and IV, and from 'very much' (6) to 'not at all (1) in Part V" (Gasmi, 2017, p. 77). It also showed acceptable internal consistency levels and the Cronbach's coefficient alphas was .86.

Procedure

The present study aims to examine the relationship between Iranian EFL learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy based on quantitative design. The data collection for this study took place in May 2021.

After getting permission from the supervisors and teachers of the English institutions, the researchers distributed the questionnaires online among English language learners. The language of the questionnaires was English. To make sure that students understand the items properly, the Persian translation of the items was sent to them too. Learners completed the scales in two hours. After gathering the questionnaires, the data were analyzed.

Data Analysis

The data were analyzed by the SEM approach of Amos software to find more exact correlations among the variables. The data were also analyzed by SPSS software for Descriptive statistics of variables of the study, Reliability, and Pearson correlation. In addition, the proposed model was tested using the Amos 24 statistical package.

RESULTS

Descriptive statistics of variables of the study (learners' flipped learning readiness, learning engagement, critical thinking, and autonomy), are presented in Table 1.

Table 2 summarizes the information obtained from Cronbach alpha analyses. As can be seen, the utilized questionnaires gained acceptable indices of Cronbach alpha

The alpha coefficient for total Flipped Learning Readiness with 27 items (.89), for total Learning Engagement with 65 items (.91), for total Critical Thinking with 34 items (.73), and total Autonomy with 65 items (.84), suggest that the items have relatively good internal consistency.

Table 1

Descriptive Statistics of the Variables

	N	Min	Max	Mean	SD
Flipped Learning Readiness	520	56	127	103.24	5.45
Learning Engagement	520	189	323	242.18	21.87
Critical Thinking	520	11	32	20.74	2.09
Autonomy	520	177	287	211.56	17.04

Table 2

Results of Cronbach alpha

Scale	Number of Items	Cronbach alpha
Flipped Learning Readiness	27	.89
Learning Engagement	65	.91
Critical Thinking	34	.73
Autonomy	65	.84

Q: Are there any statistically significant relationships between EFL learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy?

To assess Research Question, Pearson correlation was used. Table 3 indicates the results of the correlation between learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy.

Results of correlation revealed that total flipped learning readiness correlated positively and significantly with three variables: learning engagement (r=.46, p=.000), critical thinking (r=.25, p=.000), and autonomy (r=.39, p=.000). Based on this result, flipped learning readiness has the highest correlation with learning engagement and the lowest correlation with critical thinking. In addition, the correlation analysis indicates that while statistically significant, the strength of the correlations is overall quite weak between FLR, CT, AU, and LE.

In addition, the proposed model was tested using the Amos 24 statistical package. To check the strengths of the causal relationships among the components, standardized estimates were examined. Several fit indices were examined to evaluate the model fit. Table 4 shows the goodness of fit indices.

Table 4 shows the chi-square/df ratio (2.55), RMSEA (.07), GFI (.95), NFI (.91), and CFI (.92), all the fit indices lie within the acceptable fit thresholds. Therefore, it can be concluded that the proposed model had a perfect fit with the empirical data.

Figure 1 represents the schematic relationships between EFL learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy.

As indicated in Figure 1, flipped learning is a positive significant predictor of three variables: Critical Thinking (β= .18, p<0.05), Engagement (β= .31, p<0.05), and Autonomy (β= .24, p<0.05). In addition, engagement is positively predicted by both critical thinking (β= .16, p<0.05), and autonomy (β= .22, p<0.05). Finally, results revealed that Critical Thinking is a positive significant predictor of Autonomy (β= .17, p<.05).

DISCUSSION

One of the findings of this study is that there was a positive relationship between flipped instruction and learners' critical thinking. This finding is in line with the finding of Viriyavejakul's (2020) study. He also found that flipped instruction and critical thinking are positively interrelated. In this respect, Green (2015) argued that engaging students in the process of learning such as in flipped learning could develop their critical thinking ability as well as problem-solving skills. Also, the other finding of this study is that flipped learning readiness had the highest correlation with learning engage-

Table 3

Results of Correlation between Variables

		1. FLR	2. LE	3. CT	4. AU
1. Flipped Learning Readiness	Pearson Correlation	1			
	Sig. (2-tailed)				
	N	520			
2. Learning Engagement	Pearson Correlation	.46**	1		
	Sig. (2-tailed)	.000			
	N	520	520		
3. Critical Thinking	Pearson Correlation	.25**	.18*	1	
	Sig. (2-tailed)	.000	.011		
	N	520	520	520	
4. Autonomy	Pearson Correlation	.39**	.35**	.22**	1
	Sig. (2-tailed)	.000	.000	.001	
	N	520	520	520	520

***. Correlation is significant at the 0.01 level (2-tailed).*
**. Correlation is significant at the 0.05 level (2-tailed).*

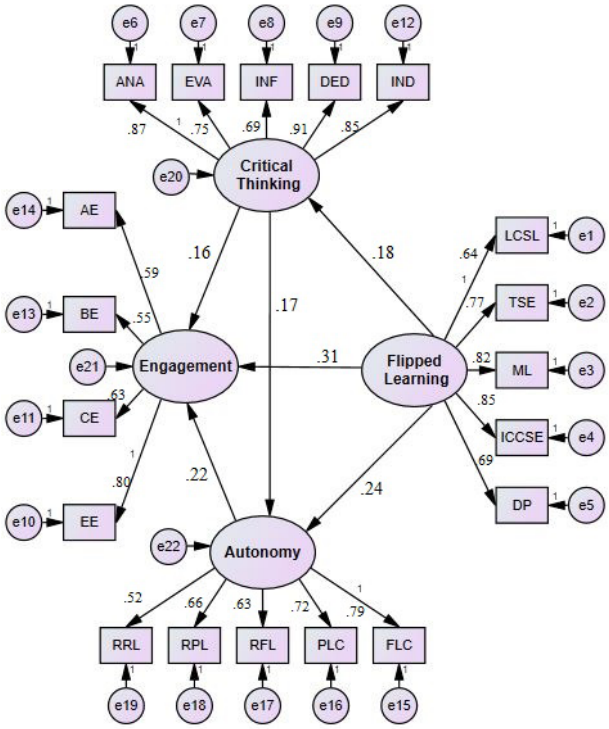
Table 4

The goodness of fit indices

	X2/df	GFI	CFI	NFI	RMSEA
Acceptable fit	<3	>.90	>.90	>.90	<.08
Model	2.55	.95	.92	.91	.07

Figure 1

The schematic relationships among variables



Note. AE: agentic engagement, BE: behavioral engagement, CE: cognitive engagement, EE: emotional engagement, ANA: analysis, EVA: evaluation, INF: inference, DED: deduction, IND: induction, LCSL: learner control and self-directed learning, TSE: technology self-efficacy, ML: motivation for learning, ICCSE: in-class communication self-efficacy, DP: doing previews, RRL: recognition of responsibility for learning, RPL: responsibility for past learning, RFL: responsibility for future learning, PLC: past learning outside the classroom, and FLC: future learning outside the classroom.

ment. The result of this study confirms the results of Gasmi's (2017) research that found that flipped instruction had an impact on EFL learners' engagement in academic writing classes in Oman. This may be because, in flipped learning, students become more actively involved with learning.

In addition, the other finding of this study is that there was a positive relationship between flipped instruction and learners' autonomy. The result of this study is in line with the result of Han (2015) who found that flipped classrooms had a significant effect on ESL learners` autonomy. As flipped learning needs students to be actively engaged in learning, the development of learner autonomy could be achieved. In this form of instruction, because students are not greatly dependent on their teachers in the process of learning and teaching such as in the traditional ones, they get more autonomy. When students are at higher levels on all the flipped learning readiness dimensions, they will be more independent and autonomous. This may be because of "successful online learning experiences leading them to be more familiar with the use of information technology for learning and also with self-directing their learning" (Hao, 2016). Thus, these students are more confident (Ferlazzo, 2015) and autonomous, motivating them more to do previews. Therefore, students who have a better-flipped learning readiness may tend to be more able of and be more motivated to learning engagement, to be autonomous in learning and to do self-directed learning, and have higher opportunities for critical thinking in their process of learning.

Many implications arise from this study. Students are regarded as the first beneficiary of the results. The relationships of readiness levels, learning engagement, critical thinking, and autonomy show that enhancing students' readiness is essential in improving their learning factors. Being taught through flipped instruction, students can develop their learning engagement, critical thinking, and autonomy because they are exposed to an interactive and reflective learning context in which they appear to be more interested to improve their learning opportunities by finding their way of learning. When students are ready for flipped instruction, they can take the required action to solve the probable problems in their learning. Second, teachers need to assure learners of their abilities, competencies and indicate empathy with their learners to develop readiness levels. Technology-based instruction by flipped one can help teachers to achieve their aims by engaging their students in learning, which makes it more enjoyable, interesting, and meaningful for the students. Also, teachers may find it beneficial to instruct on how to take advantage of online learn-

ing resources and to improve their learning experience. In addition, teacher trainers also require developing teacher training programs and workshops to increase the teachers' awareness of teaching and managing their flipped instruction classes.

Finally, the researcher hopes the findings would shed light on more influential instructional design strategies for the design and implementation of the flipped instruction approach according to learners' perspectives. The results of this study also would develop a framework for the instructional designers, teachers, teacher trainers, institute managers, and supervisors to develop and perform flipped instruction according to approved instructional standards and principles.

The significant limitation of this study is that learners' flipped learning readiness, learning engagement, critical thinking, and autonomy were evaluated by the questionnaires. No interviews or open-ended questions were included so that the students could voice their own thoughts. Further studies can do a mixed-method approach and evaluate them by both questionnaires and interviews or open-ended questions. Also, in this study, the researcher examined the relationship between learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy among language institute learners. Further research can examine the role of these variables in public schools and compare them with the results of the present study. Also, this research should be repeated with more learners from different parts of the world and use methods that guarantee a higher level of randomization and at last greater generalizability. Also, other research can be implemented to explore the relationship between learners' flipped learning readiness with other variables such as students' motivation and achievement.

CONCLUSION

As stated before, the present study sought to examine the relationship between Iranian EFL learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy. With this aim, a research question was raised: "Are there any statistically significant relationships between EFL learners' flipped learning readiness and their learning engagement, critical thinking, and autonomy?" To get a clear picture of the yield results, the research question was addressed by examining the proposed model with SEM and using the Pearson correlation coefficient. The results of the research question revealed that total flipped learning readiness correlated positively and significantly with three variables: learning engagement, critical thinking, and autonomy. Based on this result, flipped learning readiness had the highest correlation with learning engagement and the lowest correlation with critical thinking. Also, the proposed model had a perfect fit with the empirical data. According to the findings, flipped learning was a positive significant predictor of three variables: critical thinking, engagement, and autonomy. In addition, engagement was positively predicted by both critical thinking, and autonomy. Finally, results revealed that critical thinking was a positive significant predictor of autonomy.

DECLARATION OF COMPETING INTEREST

None declared.

REFERENCES

Arum, R., & Roska, J. (2011). *Academically adrift: Limited learning on college campuses.* http://www.press.uchicago.edu/ucp/books/book/chicago.A/bo10327226.HTML

Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education.

Berrett, D. (2012). How 'flipping the classroom can improve the traditional lecture. *The Chronicle of Higher Education, 12*(19), 1-3.

Bishop, J. L., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. In *ASEE National Conference and Exposition* (Paper ID 6219). Atlant.

Engin, M., & Donanci, S. (2016). Instructional videos as part of a 'flipped' approach in academic writing. *Learning and Teaching in Higher Education: Gulf Perspectives, 13*(1), 1–8. https://doi.org/10.18538/lthe.v13.n1.231

Ferlazzo, L. (2015). *Building a community of self-motivated learners: Strategies to help students thrive in school and beyond*. Routledge

Gasmi, A. A. (2017). *An investigation of the impact of flipped instruction on EFL students' engagement in academic writing classes: A case study of foundation students in Oman* [Unpublished doctoral dissertation]. University of Liverpool. http:// H00022808_Nov2017.pdf

Graham, C. R., Woodfield, W., & Harrison, J. B. (2013). A framework for institutional adoption and implementation of blended learning in higher education. *The Internet and Higher Education, 18*, 4-14. https://doi.org/10.1016/j.iheduc.2012.09.003

Green, T. (2015). Flipped classrooms: An agenda for innovative marketing education in the digital era. *Marketing Education Review, 25*(3), 179-191. https://doi.org/10.1080/10528008.2015.1044851

Han, Y. J. (2015). Successfully flipping the ESL classroom for learner autonomy. *NYS TESOL Journal, 2*(1), 98-109. https://doi.org/110481685

Hao, Y. (2016). Middle school students' flipped learning readiness in foreign language classrooms: Exploring its relationship with personal characteristics and individual circumstances. *Computers in Human Behavior, 59*, 295-303. https://doi.org/10.1016/j.chb.2016.01.031

Herreid, C. F., & Schiller, N. A. (2013). Case studies and the flipped classroom. *Journal of College Science Teaching, 42*(5), 62-66.

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2015). *NMC Horizon Report: 2015 K-12 edition.* The New Media Consortium.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education, 31*(1), 30-43. http://dx.doi.org/10.2307/1183338

Lee, H., & Wang, P. (2013). EFL college student perceptions, engagement, and writing developments in a wiki-based inter-university collaborative writing project. *English Teaching & Learning, 37*(2), 77-120. https://doi.org/10.6330/ETL.2013.37.2.03.

Mason, G. S., Shuman, T. R., & Cook, K. E. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Transactions on Education, 56*(4), 430-435. https://doi.org/10.1109/TE.2013.2249066

McLaughlin, J. E., Roth, M. T., Glatt, D. M., Gharkholonarehe, N., Davidson, C. A., Griffin, L. M., Mumper, R. J. (2014). The flipped classroom: a course redesign to foster learning and engagement in a health professions school. *Academic medicine, 89*(2), 236-243. https://doi.org/10.1097/ACM.0000000000000086.

Miserandino, M. (1996). Children who do well in school: Individual differences in perceived competence and autonomy in above-average children. *Journal of Educational Psychology, 88*(2), 203-214. https://doi.org/10.1037/0022-0663.88.2.203

Moranski, K., & Kim, F. (2016). "Flipping" lessons in a multi-section Spanish course: Implications for assigning explicit grammar instruction outside of the classroom. *The Modern Language Journal, 100*(4), 830–852. https://doi.org/10.1111/modl.12366

Muniandy, A/P., V. (2018). Effectiveness of flipped classroom on students` achievement and attitudes towards the English language in secondary school. *Journal of Innovative technologies in Education (JITE), 2*, 9-15. https://doi.org/167213939

O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *The Internet and Higher Education, 25,* 85-95. https://doi.org/10.1016/j.iheduc.2015.02.002

Öztürk, M., & Çakıroğlu, U. (2021). Flipped learning design in EFL classrooms: implementing self-regulated learning strategies to develop language skills. *Smart Learning Environments*, 8(2), 1-20. https://doi.org/10.1186/s40561-021-00146-x

Reeve, J., & Tseng, C. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology, 36*(2011), 257-267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Sadeghi, B., Hassani, M. T., & Rahmatkhah, M. (2014). The Relationship between EFL learners" metacognitive strategies, and their critical thinking. *Journal of Language Teaching and Research, 5*(5), 1167-1175. https://doi.org/10.4304/jltr.5.5.1167-1175

Sakai, S., Chu, M., Takagi, A., & Lee, S. (2008). Teachers' roles in developing learner autonomy in the East Asian Region. *The Journal of ASIA TEFL, 5*(1), 93–117. https://doi.org/10.12973/iji.2017.1018a

Sezer, R. (2008). Integration of critical thinking skills into elementary school teacher education courses in Mathematics. *Education, 128*(3), 349-362.

Shahani, S., Chalak, A., & Heidari Tabrizi, H. (2021). The Iranian intermediate EFL learners` attitudes towards using flipped teaching via Google Classroom. *Research in English Language Pedagogy, 9*(1), 90-112. https://dx.doi.org/10.30486/relp.2020.1880561.1159

Smith, P. J. (2005). Learning preferences and readiness for online learning. *Educational Psychology, 25*(1), 3-12. https://doi.org/10.1080/0144341042000294868

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation, and task orientation. *Learning Environments Research, 15*(2), 171-193. http://dx.doi.org/10.1007/s10984-012-9108-4

Touchton, M. (2015). Flipping the classroom and student performance in advanced statistics: Evidence from a quasi-experiment. *Journal of Political Science Education, 11*(1), 28-44. https://doi.org/10.1080/15512169.2014.

Tecedor, M., & Perez, A. (2019). Perspectives on flipped L2 classes: implications for learner training. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2019.1626439

Thorndike, E. (1932). *The fundamentals of learning*. ASS Press.

Vaezi, R., Afghari, A., & Lotfi, A. (2019). Flipped teaching: Iranian students' and teachers' perceptions. *Applied Research on the English Language, 8*(1), 139-164. https://dx.doi.org/10.22108/are.2019.114131.1382

Viriyavejakul, C. (2020). The effect of flipped education on undergraduate student critical thinking ability. *International Journal of Online Pedagogy and Course Design*, 10(4), 1-9. https://doi.org/10.4018/IJOPCD.2020100105

Yarbro, J., Arfstrom, K. M., McKnight, K., & McKnight, P. (2014). *Extension of a review of flipped learning*. Flipped Learning Network.

https://doi.org/10.17323/jle.2022.12935

Cognitive Predictors of Coherence in Adult ESL Learners' Writing

Abdul Saeed [1,2] , John Everatt [2] , Amir Sadeghi [3] , Athar Munir [2,4]

[1] Sukkur Institute of Business Administration University, Pakistan

[2] University of Canterbury, New Zealand

[3] PEETO, Multicultural Learning Centre, New Zealand

[4] Emerson University Multan, Pakistan

ABSTRACT

Background. Coherence is considered one of the most important qualities of written discourse. Despite its fundamental importance, it is still considered a fuzzy and abstract concept in most English Second Language (ESL) contexts. Consequently, many ESL learners struggle to produce a coherent text. Morphological, phonological, orthographic awareness, vocabulary knowledge, and grammatical competence have been identified as predictors of writing quality in novice writers. There is, however, a lack of data to assess whether such linguistic skills also predict coherence in adult ESL learners' writing.

Purpose. The purpose of the study was to find out the relationships among a set of linguistic skills measures which included morphological, phonological, orthographic awareness, vocabulary knowledge and grammatical competence and coherence in adult ESL learners' writing.

Methods. To testify to the potential predictors of coherence in ESL writing, adult university students (126) were assessed by the measures of the linguistic skills mentioned above in addition to four measures of coherence: two relatively reader-based measures (ILETS and the Holistic Coherence Scale) and two relatively text-based measures (Topical Structure Analysis and Topic Based Analysis). All measures of the study were proved valid and reliable.

Results. The findings revealed that vocabulary knowledge, morphological awareness, and grammatical competence were related to the coherence measures, particularly the reader-based measures. In contrast, measures of phonological and orthographic awareness generally did not correlate with the coherence measures.

Implication. Reasons for the associations among the variables of the study were discussed and areas for future research were offered.

KEYWORDS

Coherence, second language writing, predictors of writing quality

INTRODUCTION

Writing is considered a more complex and difficult skill to learn than listening, speaking, and reading. It is because writing does not come naturally like listening and speaking. To have good writing skills, one must not only learn them but also practise them regularly. Writing is very close to reading in this regard. The situation gets further complex for writing in a second language (L2) as the learner has to learn a new set of writing skills that may be different from their first language. Faced with the complexities of this process, and to support learners and make the writing process easier both in L1 and L2, researchers have attempted to identify predictors of writing. High proficiency in linguistic skills seems to predict writing quality. Such predictors include, but are not limited to, morphological, phonological, and orthographic awareness, vocabulary knowledge and grammatical competence. Studies of writing predictors have addressed such skills among writers (for example, Berninger et al., 2010; McCutchen et al., 2014; Abu-Rabia (2001); see literature review for further details on studies of ESL students) but few studies have considered predictors of measures of coherence among adult ESL learners.

Citation: Saeed, A., Everatt, J., Sadeghi, A., & Munir, A. (2022). Cognitive Predictors of Coherence in Adult ESL Learners' Writing. *Journal of Language and Education, 8*(3), 106-118. https://jle.hse.ru/article/view/12935

Correspondence:
Abdul Saeed,
saeedabdulskr@gmail.com

Received: August 26, 2021
Accepted: September 18, 2022
Published: September 30, 2022

Coherence is an essential and important construct to assess one's quality of writing (Candelo et al., 2018; Chiang, 2003). Traditionally, coherence has been defined as the semantic relationship in the text whereby all elements are logically joined to give a single unit of meaning (Knoch, 2007). Yet, in spite of its fundamental importance in writing, it is deemed to be a fuzzy and abstract concept in most ESL contexts (Lee, 2002), a misconception that often leads to its neglect in teaching and learning (Attelisi, 2012). One consequence of this deficit is that many adult ESL learners struggle to produce coherent texts (Masadeh, 2019; RahmtAllah, 2020). In addition, coherence has been considered a subjective construct (Van Dijk, 1977). For this reason, most measures of coherence have been seen as subjective and it has proven difficult to find a completely objective measure of coherence analysis (Todd, 2016).

The purpose of the present study was to investigate a set of linguistic skills (i.e., morphological, phonological, and orthographic awareness, vocabulary knowledge and grammatical competence) as potential predictors of measures of coherence in adult ESL learners' writing. The findings should inform theories about coherence analysis and help language teachers and ESL learners to focus on key skills involved in writing a coherent text. The current research aimed to answer the following questions:

(1) Are there any relationships among linguistic skills such as morphological awareness, phonological awareness, orthographic awareness, vocabulary knowledge and grammatical competence and coherence in adult ESL learners' writing? If there are, what are the best or common predictors of coherence?

(2) Are there any relationships among the sub-component parts of the measures of coherence used in this study and the measures of morphological awareness, phonological awareness, orthographic awareness, vocabulary knowledge and grammatical competence?

LITERATURE REVIEW

Coherence

Coherence is deemed to be one of the most important qualities of writing (Attelisi, 2012; Crossley & McNamara, 2010). It is known as 'sine qua non' (a thing that is necessary) in written discourse (McCulley, 1985). Its centrality can also be gauged by the fact that it is present in every test of English in which learners' proficiency is assessed, such as the Test of English as a Foreign Language (TOEFL), and the International English Language Testing System (IELTS). Nevertheless, in its centrality in written language learning, coherence remains a fuzzy and abstract concept to most ESL and EFL teachers, many of whom are unable to define and practise coherence in their classrooms (Lee, 2002). As a consequence, a great number of ESL/EFL learners struggle to produce highly coherent texts (Khalil, 1989).

Predictors of Writing Quality

The current research investigated potential predictors of coherence in ESL writings. A range of linguistic skills that have been identified to predict variance in literacy skills (reading and writing) were further examined in the current study; the predictors that are mostly suggested in works focused on children. This is because such skills are considered to be developing in children but may have fully developed in adults. However, the development of such skills may vary across first language and second language contexts. It can be argued that adult ESL learners may still develop such skills in a second language which can influence their ability to produce a coherent text. This background review of the literature, therefore, focuses on the evidence of linguistic skills targeted in the study that may predict literacy outcomes – previously suggested in studies among children.

Morphological Awareness

Morphological awareness, an important element of language learning, is generally known as a conscious awareness of the morphemic structures of the words and the ability to manipulate that structure to make new words. Zhang and Koda (2012) conducted a study to test the role of morphological awareness in the development of second language vocabulary and reading comprehension amongst 130 students in a university in Shanghai in China. The study concluded that morphological awareness contributed significantly to the development of vocabulary knowledge in English as a second language. In addition to this, morphological awareness has also been found to have positive effects on learning spelling in novice writers, particularly in the spelling of morphologically complex words (Berninger et al., 2010; McCutchen and Stull, 2015).

Phonological Awareness

Phonological awareness has also been documented as a predictor of writing in children. Mackenzie and Hemmings (2014) examined the role of phonemic awareness in the development of English writing performance of children in ten kindergarten classrooms in New South Wales, Australia. Findings indicated a high correlation between oral language performance and writing development in these young children. Additionally, Zhao (2011) tested 339 grade 8 students in China with Chinese first and English as a second language. The participants were measured on their morphological, phonological, and orthographic awareness skills. Results revealed orthographic awareness as the main contributor to spelling in Chinese (the first language), whereas phonological awareness was the main predictor for spelling in English.

Orthographic Awareness

Abu-Rabia (2001) found orthographic awareness a predictor of spelling ability in both English and Russian languages. The participants were ESL university students in Israel between the ages of 25 and 30. Findings suggested that orthographic awareness may improve learners' spelling which can further develop their writing skills since improved spelling may help writing. Harrison et al. (2016) conducted research to assess potential predictors of spelling and writing in grade 3 in native and ESL learners. The participants were given rapid naming, phonological awareness, single-word fluency, text spelling, handwriting fluency, and paragraph writing fluency tasks. For native speakers, vocabulary and rapid naming predicted variance in writing whereas rapid naming and syntactic awareness predicted writing quality in ESL learners. Orthographic awareness was found a predictor of spelling in native as well as ESL learners.

Vocabulary Knowledge

Diamond et al. (2008) described vocabulary knowledge as the comprehension of the meaning of words in different contexts. To convey a message effectively, a learner needs to have good vocabulary knowledge. In other words, poor text generation and comprehension could both be caused by a lack of vocabulary knowledge (Lee, 2003). To examine the relationship between lexical diversity, indicating vocabulary knowledge, and holistic markings of a composition, Roessingh et al. (2015) asked 77 third-grade students in Canada to write a composition on a given prompt. Lexical diversity was assessed by corpus-based analysis through high-and-low-frequency words. For holistic scoring, a trait-based rubric (HALT) was used. The findings highlighted a correlation between lexical diversity, which is indicative of vocabulary size and writing quality.

Grammatical Competence

Grammar knowledge is believed to give the learner a sense of correct and incorrect use of the language (Wang et al., 2015). Grammar teaching has always been an inseparable part of second language teaching. There is a large body of research on different aspects of grammar teaching and their effects on language learning in both first and second language contexts (Berninger et al., 2011; Kim et al., 2013; McCutchen & Stull, 2015; O'brien et al., 2006; Wong, 2012). As an example, Jones et al. (2013) argued on the bases of findings from several schools in England that the teaching of grammar improves learners' understanding of the writing system. Their data indicated a positive effect of grammar teaching on writing, with skilled writers benefitting more than less skilled writers. Such studies support the view that higher levels of grammar knowledge should predict writing skills.

METHODS

Participants

Participants were recruited from six different government universities in Punjab, Pakistan. Following informed consent, 129 university students volunteered. Three participants were excluded from the data. Two students did not complete the measures: one missed a session due to illness and another had to leave the test session due to a personal emergency. The third case was the student who scored zero in the English vocabulary measure. Out of the remaining 126, 58 were male and 68 were female, with their ages ranging from 18 to 27 years (M = 21 years).

Background information was provided by the participants via a demographic questionnaire. All students were multilingual speakers, with the majority (82) being Punjabi speakers, though a minority used Saraiki as their first language. They all were able to communicate in Urdu, the national language of Pakistan, as well as in English, and for most, these were also the languages of reading – a minority could also read in their home language.

All participants have learned English since their first year of schooling in Pakistan. Additionally, the participants reported to have been using English in verbal communication between 5 and 22 years (M= 14 years). All participants were university students enrolled in their second year of study with English being the language of instruction according to the standards set by the [1]Pakistan National Curriculum (Ministry of Education, 2006). The National Curriculum for English Language grades I-XII (Ministry of Education, 2006, pp. 127-131) requires students to be able to display all necessary writing skills, such as brain storming, mind mapping, transitional devices, key ideas, grammar, syntactic maturity and variety of sentences. Hence, the participants were assumed to know how to produce a well-connected coherent text. Additionally, since the participants were university students in their second year, they also experienced writing their assignments in English at tertiary levels for at least one year. A majority of these students would fall under the B1 category of the Common European Framework of Reference for Languages, which describes independent users of English who can use language for a variety of purposes, read extended texts, use correct and varied sentence structures, and also produce well-connected text. The criterion established by the Higher Education Commission for students with certain academic standards, as was described earlier in this paragraph, can serve as further validation. As these students are in the second year of their degree and have done with their first year, it is understood that a majority of participants have B1level of proficiency.

[1] Ministry of Education. (2006). *National curriculum for English language grades I-XII. Government of Pakistan*. https://bisep.com.pk/downloads/curriculum/Grades-I XII/pk_al_eng_2006_eng.pdf

Data Collection

To identify relationships between these language skills and measures of coherence, the participants were asked to write an essay of 250-300 words on a given topic. The topic was given via a short 150-word story about 'wildlife' and individuals who study wildlife. The participants read the story before the writing task, and they were given an hour to complete the task. They were also told to use this time in whatever way they chose (planning, writing, revising). The students were told that the essays would be assessed on several factors including meaningfulness and appropriate use of English. The topic was selected as one that participants should be at least partially familiar with. The purpose of the story was to give all participants the same topic and range of ideas for completing their essays, thereby avoiding major variations in essays due to topic knowledge and/or choice. Essays were handwritten on a sheet of paper and then transcribed onto a computer for assessment. The transcript was an exact copy of the original handwritten essay to allow an analysis of grammatical errors, organisation, and content, but the computer copy made it easier for two raters to assess the essays using the four measures of coherence.

Measures and Procedures

All measures were developed for the current study. The measures were piloted with participants from a similar background to those participating in the original study. Two or three sessions were used to administer the measures. The majority of the data were gathered over the course of two sessions; however, a third session may have been necessary if students were preoccupied with other tasks or had a different schedule. Language tests and a background questionnaire were completed during the first session. The participant's background questionnaire was given by the researcher, who also offered assistance as needed. The language tests were also given to the participants in the same session. The writing assignment was completed by the participants in the second session. To ensure the adequate distance between participants, tests were conducted in lecture halls with groups of no more than 20 students.

Language Measures

The morphological processing measure aimed to assess the participant's understanding of how words are broken down into smaller meaningful units, including roots, prefixes, and suffixes, and how words can be derived from root forms. In the current study, participants were given two tasks. Task one required the participants to choose the correct form to complete a sentence. For example, 'Geography involves the study of different (country)' and 'I (start) my new school last week' with the task being to write the correct form of the word in brackets that completed the sentence; 'countries' and 'started' in these examples. There were 20 items in this task. The second task required the participants to write a word based on the rule given in the example. For example, 'sing – singer; read – ….', where the correct answer would be 'reader'; or 'boy – boys; man – …' where the correct answer would be 'men'. There were 20 items in this task, too. Examples were provided to show what was required and each item was given one mark for a correct answer, making a potential maximum score of 40. The mean for the measure was 22.06, with a standard deviation (SD) of 8.06, and a range from 6 to 36. Cronbach's alpha was around .9.

The phonological awareness measure assessed the student's ability to use the sound pattern of the language and translate a written form into that sound pattern. Participants were given pairs of made-up words (pseudo-words) and were asked to choose the item that sounded like a real English word. For example, 'nale pult' and 'warg dore' were presented; the participants should select 'nale' and 'dore' as these sound like the English words 'nail' and 'door'. There were 17 pairs of pseudo-words in the test and each correct answer was given one mark. The mean for this measure was 12.02, SD = 2.79, and a range from 2 to 17. Cronbach's alpha was over .7.

An orthographic choice task was used to assess orthographic awareness. This task focused on the ability to identify correct spellings based on their orthographic features. The task comprised 18 pairs of letter strings, one of which was an incorrect spelling while the other was correct: for example, 'monk, munk' or 'goat, gote'. Items were selected so that both produced the same word-sound if converted by simple English letter-sound conversion rules to require recalling the orthographic features of the word to choose the correct item. The Participants were asked to choose the correct spelling in each pair and they were given one mark for each correct answer. The mean for the measure was 8.87, SD = 2.78, and a range from 2 to 16. Cronbach's alpha was over .7.

A vocabulary test was used to assess the participant's knowledge of words. In the present study, the participants were given 40 words followed by one possible meaning and three distractors for each word. The participants were asked to choose the closest meaning to the given word. For example, if the word 'rich', followed by '(i) no money at all, (ii) have a lot of money, (iii) feel happy, and (iv) feel sad', with the participant being expected to choose item ii as the correct answer. Each correct answer scored one mark. The mean for this measure was 23.27, SD = 5.85, and a range from 9 to 37. Cronbach's alpha was over .8.

Grammatical knowledge was assessed via 22 items that required the participant to identify the correct form of grammar, use of articles, and subject-verb agreement. There were three tasks in this measure to assess different aspects of grammatical understanding. In the first task, the participants were presented short sentences with four parts of the sentence underlined, one of which comprised an error based on its context. For example, in the sentences 'I am

going to an Indian restaurant for a lunch. Will you go with me? It's not too far away. It serves the best food, I believe.', the participants were required to indicate which of the underlined sections was incorrect ('a lunch' in this case). In the second task, sentences were again used, but this time with gaps, and the participants were asked to indicate one word/phrase from four options that completed the sentence. For example, the sentence might be 'The distinct geology of the island began _____ about 20,000 years ago', for which the options would be 'i) formed; ii) form; iii) to form; iv) was forming'. A final task involved a passage of text that contained grammatical errors which the participant was expected to correct. For example, for the sentences 'I could see the water from my window. the boat sailed over the waves silent.', the participant should change 'the' to 'The' and 'silent' to 'silently'. Each correct answer was given a mark. The mean for the measure was 8.37, SD = 3.50, and a range from 1.5 to 18. Cronbach's alpha was .7 or above.

Coherence Measures

For the assessment of coherence, four measures of coherence were used. These include (i) the part of the [2]International English Language Testing System (IELTS) measure that analyses coherence (IELTS, 2019), (ii) the Holistic Coherence Scale (HCS) developed by Bamberg (1984), (iii) the Topical Structure Analysis (TSA) developed by Lautamatti (1978), and (iv) the Topic Based Analysis developed by Todd (2016). To ensure consistency, two raters marked the essays. The first rater was the first author and the second was recruited as they were an experienced ESL teacher who had approximately ten years of experience in teaching and assessment. For each essay, the raters followed the criteria described below for each coherence assessment to calculate a mark – the raters made themselves familiar with the assessment procedures prior to calculating marks to ensure they understood the methods required. Correlations between the two raters' scores were .75 for the IELTS scores, .82 for the Holistic Coherence Scale, .81 for the Topical Structure Analysis, and .71 for the Topic Based Analysis. Given a reasonable level of agreement, the first rater's scores were used in the analyses as they were more familiar with the coherence methods used in the study.

The four coherence measures were selected because of the evidence for their validity and reliability provided by researchers over the years (for IELTS, see Moore, 2007; Müller & Daller, 2019; Schoepp, 2018 – for the Holistic Coherence Scale, see Connor & Lauer, 1985; McKenna, 1988 – for Topical Structure Analysis, see Ghazanfari et al., 2011; Kılıç et al., 2016; Knoch, 2007 – for Topic Based Analysis, see Todd, 2016). These measures of coherence provided a mixture of relatively old and new measures indicating the development of the topic over years: the Holistic Coherence Scale and Topical Structure Analysis were developed in the 1970s and 1980s, whereas Topic Based Analysis and the coherence scale of IELTS were comparatively new. The four measures also provided a combination of text-based and reader-based measures of coherence. The IELTS and Holistic Coherence Scale measures focused more on reader-based perspectives, whereas the Topical Structure Analysis and Topic-Based Analysis focused more on text-based methods of assessing coherence.

The International English Language Testing System (IELTS) is a high-stake English language proficiency test for international students and migrants (Alsagoafi, 2013). Candidates' writing skill is assessed on four criteria: 'Task Achievement (in Task 1) and Task Response (in Task 2), Coherence and Cohesion, Lexical Resource, and Grammatical Range and Accuracy' (Pearson, 2018). Marks are given to the writing task on a 9-point scale; (1 is the lowest level for poor outputs and 9 is the highest band – a score of 0 was not used as this indicates no attempt). Given the focus of the current study, only the Coherence and Cohesion component was used. This analysis of coherence included an assessment of cohesion, which has been considered one of the constructs of coherence (Halliday & Hasan, 2014); though equally, coherence is not simply the use of cohesive devices, and evidence has shown that a text can be coherent with a few cohesive ties (Kim & Crossley, 2018). Hence, the coherence assessment also included elements of organisation, or the compositional structure of a text (an appropriate beginning, middle and ending), and the progression or development of ideas via the logical insertion of new information related to the main topic. A score of 9 was given to texts that logically organised information and ideas, with evidence of a clear progression throughout, that used cohesion well and showed good use of paragraphing. A lower score of about 6 was given to texts that showed overall progression of ideas and information and used cohesive devices effectively, though these may be faulty/mechanical with paragraphs that were not always logically connected. A score of 4 was given to texts that presented ideas and information but in which these were not well arranged and a lack of clear progression and basic use of cohesive devices, and paragraphing may be confusing. The lowest mark of 0 was given to those who failed to convey any message. The assessment led to the participants achieving a mean score of 4.77, SD = 1.30, and a range from 1 to 8.

Bamberg (1984) developed the Holistic Coherence Scale to assess coherence for a large number of essays. This used a 4-point rubric scale with a score of 4 indicating the highest level of coherence termed as fully coherent. This suggests that the writer identifies the topic with no shift or digress, orients the reader by creating a context or situation, organises details according to a discernible plan that is sustained throughout the essay, skilfully uses cohesive ties such as lexical cohesion, conjunction, reference, etc. to link sentenc-

[2] From IELTS. (2019). IELTS scoring in detail. https://www.ielts.org/-/media/pdfs/writing-band-descriptors-task-2.ashx?la=en.

es and/or paragraphs together, and often concludes with a statement that gives closure. Additionally, the discourse flows smoothly with few or no grammatical and/or mechanical errors interrupting the reading process. Lower scores, however, indicate less coherence, with 3 suggesting partial coherence and 2 suggesting that the text is incoherent; that is, the reader can unlikely infer the topic, the writer digresses frequently and provides little orientation or organisational plan, the writer uses few cohesive ties to link text and there is no sense of closure with reading being interrupted by frequent mechanical and grammatical errors. A score of 1 indicates that the text is incomprehensible. Connor and Lauer (1985) argued that the descriptors of the scale could be divided into six sub-components of coherence: Focus, Context, Organisation, Cohesion, Closure and Grammar error. Each of these sub-components is also marked on a 4-point scale, with 4 being the highest score and 1 the lowest. Both the overall coherence scale of Bamberg and the six sub-components were calculated for each essay in the present study, though the raters started with the components, before giving an overall mark for the essay. Based on the description of the components, Focus in this context means there should be no irrelevant topic in the text: sentences should be developed logically or sequentially one after the other with no abrupt changes in topic. Context refers to a social, physical or psychological setting, and this should be clear throughout the text. Organisation suggests that the writer should start with the focus of the text and move towards a clear end. The Closure should typically involve a reiteration of the writer's purpose: the writer normally concludes the argument developed throughout the body of the text. Cohesion should be used to connect the text grammatically: i.e., the use of lexical ties such as reference, substitution, ellipsis, conjunction, collocation and lexical reiteration to connect text together as a cohesive unit. In the case of Grammar, the writer should avoid mistakes in tenses, subject-verb agreement and punctuation. Marks from all sub-components also showed reasonable correlations (r = .7 or greater) between the raters. The assessment led to the participants being marked with a mean score of 2.63, SD = 0.64 and a range from 1 to 4.

Lautamatti (1978) proposed Topical Structure Analysis to analyse coherence at the sentence, paragraph and discourse level. It was developed to examine how topics repeat, shift, and return to earlier topics in a discourse to maintain coherence. As part of the analysis, determining topics and progressions between topics is vital. For present purposes, texts were divided into t-units, or the 'minimal terminable unit', defined as an independent clause with all its dependent clauses (Hewings & North, 2006). The analysis used t-units as a measurement unit because of their flexibility to recognise more than one topic in compound sentences. In addition, t-units have been used by researchers to analyse learner's writing quality (Knoch, 2007; Witte & Faigley, 1981). For t-units, the analysis then determined the t-unit topic, or theme, and identified information about the topic, or rheme. The topic was always the semantic topic of the sentence, rather than the grammatical subject of a phrase, to allow for an analysis of coherence. The theme should be information already known to the reader, while the rheme should provide new information, with the juxtaposition of old/new or known/unknown information providing the basis on which the discourse topic is developed. This development was determined via different types of progressions through the text.

Four main types of progressions were considered in the present analysis. The first were Parallel progressions, where the topics in successive sentences were either the same or synonyms and/or pronouns were used to link the topics. The second type, Sequential progressions, indicated semantically related but different topics in successive sentences. Typically, this occurred when a rheme part of a preceding sentence became the topic or theme of the following sentence. The third type was Extended parallel progressions where two semantically identical topics were interrupted by at least one occurrence of a Sequential progression. These three progressions support the thematic development of the text. Parallel progressions give depth to the topic: by repeating the same topic in consecutive sentences, the writer provides additional and detailed information about the topic under discussion. Sequential progressions provide a way to extend the text by introducing new but related topics, and Extended parallel progressions remind the reader about the main topic by repeating a previously used topic. Extended parallel progressions can also be used to present a closing statement. Thus, these three progressions work to give depth, width and to close the text. The final type used in the current analyses was Unrelated progressions in which the topic of a sentence was not related to the theme or rheme part of either the preceding or the successive sentences. These types of progressions indicate problems with the coherence of the text. Combining the differing types of progressions between t-units in the text provided a basis on which to assess coherence. (Note that correlations in determining t-units and types of progressions between raters were .9 or greater.)

The specific form of this assessment was based on the marking scale developed by Knoch (2007). According to Knoch, this five-point scale has been assessed for its validity, in comparison with professional raters, and its reliability. The scale has also been used with second language learners of English (Knoch, 2007). Scores range from 4 to 8, with 8 being the highest mark reflecting frequent Sequential progressions, infrequent but supportive Parallel progressions, few but appropriate Extended parallel progressions and no Unrelated progressions. Lower scores on the scale reflected more mixed use of progressions, particularly an increase in Unrelated progressions. The lowest score represented frequent use of Unrelated progressions and infrequent use of Sequential progressions. The assessment led to the partic-

ipants being marked with a mean score of 5.88, SD = 0.64, and a range from 4 to 8.

Topic Based Analysis was developed by Todd (2016). It analyses coherence by dividing the text into t-units, tracing out references, identifying key concepts and linking these concepts through moves in the text. Todd (2016) found a high correlation between experienced teachers' marks for coherence and the number of moves per t-units assessed via Topic-Based Analysis. The same calculation was used in the present study. As above, t-units were determined as an independent clause with all of its dependent clauses. Moves were used to show the change of concept from old to new, meaning that fewer moves would be indicative of fewer concepts in the text and, hence, greater unity in the text. To identify moves, concepts within the text needed to be determined. A concept was taken as a psychological construct and represents some entity in the world. The frequency of occurrence represented the importance of a concept in the text: higher frequency concepts were considered as the topics of the text (de Beaugrande, 1981). Once t-units, concepts within t-units and moves between concepts were each determined, coherence was calculated by dividing the total number of moves in the text by the total number of t-units. (Correlations between raters were .92 for t-units and .81 for the number of moves.) A higher score, therefore, represented more moves and less unity: higher scores meant less coherence. The assessment led to the participants being marked with a mean score of 1.40, SD = 0.33, and a range from 0.75 to 2.50.

RESULTS

The aim of the study was to investigate the potential cognitive linguistic predictors of coherence. First correlations between the language measures were calculated to ensure there was variability between the skills used for each of the measures; that is, they were not simply determined by English usage/proficiency alone. Correlations were all significant (as expected) and ranged from .34, for grammatical knowledge and phonological awareness, to .7, for grammatical knowledge and morphological processing. Morphological processing also produced a correlation with the vocabulary of .67 but was less related to phonological and orthographic awareness (.43 and .55 respectively). These correlations were consistent with common elements between the measures but indicated that they were not simply determined by proficiency. Similar correlational analyses for the coherence scales also indicated variability, with the largest being between the IELTS scores and the Holistic Coherence Scale (r = .59). There was a reasonable relationship between the Topic Based Analysis and the measures of Topical Structure Analysis (-.44) and Holistic Coherence Scale (-.45), with the other correlations ranging in size from .32 to .36. Again, these were indicative of some common construct, but variability in the way the construct was conceptualised.

Relationships between Language Skills Measures and Coherence Measures

Table 1 displays the correlations between the five language measures and four measures of coherence used in the study. The results indicated that measures of vocabulary size, morphological processing, and grammatical knowledge overall produced higher correlations than those found for the phonological and orthographic awareness measures. However, the size of the correlations varied across the four coherence measures utilised in the study. Morphological processing, vocabulary size and grammatical knowledge showed medium size correlations (i.e., around .3 to .5) with the IELTS assessment and the Holistic Coherence Scale. Morphological processing produced the largest correlation with the IELTS scores (.51) and also produced the second largest with the Holistic Coherence Scale (.46), though the correlation between the Holistic Coherence Scale and the grammatical knowledge measure was almost identical (r = .45). Scores produced via the Topical Structure Analysis and the Topic Based Analysis, however, showed small correlations (less than .3) with morphological awareness, vocabulary knowledge and grammatical competence (note that negative correlations with Topic Based Analysis are because lower scores were indicative of higher levels of coherence). The measures of phonological and orthographic awareness also showed generally small correlations (less than .3) with the coherence measures, except for the correlations with the Holistic Coherence Scale (.33 and .35); and the correlations with the Topical Structure Analysis were near-zero. These findings suggest that coherence in adult ESL learners' writing is more likely associated with morphological pro-

Table 1

Correlations between the Language Skills Measures and the Coherence Measures

	Morphological Processing	Phonological Awareness	Orthographic Awareness	Vocabulary Size	Grammatical Knowledge
IELTS	.51	.22	.27	.35	.44
Holistic Coherence Scale	.46	.33	.35	.38	.45
Topical Structure Analysis	.20	.00	-.02	.21	.19
Topic Based Analysis	-.20	-.11	-.19	-.08	-.18

Note: Correlations in bold are significant at the .01 level, those in italics are significant at the .05 level

cessing, vocabulary size and grammatical knowledge than phonological and orthographic awareness.

These correlational results were consistent with the findings of regression analyses assessing the level of prediction offered by the five language skills measures for each of the coherence analyses. The results (see Table 2) present the total variability explained by the five language measures along the standardised beta scores for each measure (significant beta scores are bolded). Total variability explained was larger for the two more reader-based assessments (27-28%) but provided little explanation for the two more text-based analyses (10% for the Topical Structure Analysis, but not significantly greater than zero for the Topic Based Analysis). The morphological processing measure showed the largest beta score, with the IELTS scores. Grammatical knowledge also showed a significant beta score with the Holistic Coherence Scale, but generally beta scores were around .2 or less.

Relationships between Language Skills Measure and Sub-components of coherence Measures

Table 3 shows the correlations between the five language measures and the sub-components of the measures of coherence used in this study (note that the IELTS score was a whole score and did not have sub-components). The results indicated that both sub-components of the Topic Based Analysis (number of moves and t-units) produced small or near-zero correlations with the language skills measures. The findings also revealed relatively small correlations (less than .3) between the sub-components of the Topical Structure Analysis and the language skills measures, with the Unrelated progression measure producing the largest correlation, that with vocabulary size (r = -.29). This suggests a trend for those with smaller vocabularies to produce more Unrelated progressions in their written text. The largest correlations were found with the sub-components of the Holistic Coherence Scale. The largest correlations were again with the measure of morphological processing; though, vocabulary size and grammatical competence also produced reasonable size of correlations with the different sub-components. Focus, Organisation and Closure showed correlations with at least one language measure above .4, whereas Context, Cohesion and Grammar errors generally showed small correlations with all the language skills measures. Overall, these correlations support the argument for associations between several sub-components of this coherence scale and more meaning-based language skills (vocabulary and morphology) along with a grammatical understanding of English. However, generally, the findings again suggest small relationships between the five language skills assessed in the current study and the different aspects of coherence assessed by the four coherence analyses performed.

DISCUSSION

The findings suggested a higher relationship between morphological awareness, vocabulary knowledge and coherence. The relationship between morphological awareness

Table 2

Results of Regression Analyses for the IELTS Scores, the Holistic Coherence Scale (HCS), the Topical Structure Analysis (TSA) and the Topic Based Analysis (TBA)

	IELTS		**HCS**		**TSA**		**TBA**	
	R²	***Sig.***	***R²***	***Sig.***	***R²***	***Sig.***	***R²***	***Sig.***
Total variability explained	.28	F=9.15 p < .001	.27	F=8.72 p < .001	.10	F=2.52 p = .03	.06	F=1.52 p = .19
	Beta		Beta		Beta		Beta	
Morphological Processing	.41	t=3.14 p=.002	.18	t=1.37 p=.17	.12	t=0.84 p=.40	-.17	t=-1.14 p=.26
Phonological Awareness	.02	t=0.16 p=.87	.12	t=1.24 p=.22	-.07	t=-0.69 p=.49	-.00	t=-.03 p=.97
Orthographic Awareness	-.05	t=-0.49 p=.63	.05	t=0.50 p=.62	-.21	t=-1.84 p=.07	-.13	t=-1.12 p=.27
Vocabulary Size	.01	t=.05 p=.96	.08	t=0.68 p=.50	.20	t=1.61 p=.11	.13	t=1.07 p=.29
Grammatical Knowledge	.17	t=1.51 p=.13	.22	t=2.00 p=.05	.13	t=1.07 p=.29	-.07	t=-0.52 p=.61

Note: Collinearity statistics suggested no problems with multicollinearity in any of the analyses (i.e., tolerance scores were 0.35 or greater and VIF scores were all less than 3)

Table 3

Correlations between the Language Skills measures and the Sub-components of Coherence Measures

	Morphological Processing	Phonological Awareness	Orthographic Awareness
Focus	.51	.33	.33
Context	.33	.26	.24
Organisation	.49	.30	.21
Closure	.41	.21	.19
Cohesion	.34	.18	.23
Grammar	-.33	-.10	-.19
Parallel progressions	.23	.18	.23
Sequential progressions	.23	.05	.01
Extended Parallel progressions	-.10	-.10	-.09
Unrelated progressions	-.22	-.11	-.05
Number of t-units	.05	.02	.11
Number of moves	-.05	-.05	.00
Number of words	.26	.15	.20

Note: Correlations in bold are significant at the .01 level, those in italics are significant at the .05 level

and the development of coherence in adult ESL learners' writing may be because both are involved in the processing or production of meaning – and perhaps the awareness of how to process or produce meaning from different units of meaning: affixes and roots versus concepts, phrases, sentences, and paragraphs. Some studies that have addressed the role of morphological awareness in the development of spelling (Berninger et al., 2010; McCutchen et al., 2014), vocabulary knowledge (Mochizuki & Aizawa, 2000), and text generation (McCutchen & Stull, 2015; Northey et al., 2016). One potential relationship here is that morphologically constructed words, such as 'firstly' and 'secondly', may help the writers to organise the development of text from sentence to paragraph, and forward to the discourse level.

Vocabulary knowledge was also found to be a small but potentially significant predictor of different aspects of coherence, after morphological awareness. Vocabulary knowledge is helpful for learners to express a message using a variety of words. Like morphological awareness, vocabulary knowledge should also help learners produce a meaningful text. Thus, both coherence and vocabulary are meaning-related, and this may be the potential reason to have a significant correlation between them. These findings also seem consistent with the correlation between Sequential progression and vocabulary knowledge. For Sequential progression, writers should possess a good range of vocabulary in order to describe the same topic but using new words. It is worth noting that the correlation between Sequential progressions and morphological awareness is similar to that between Sequential progressions and vocabulary knowledge and, therefore, there may be alternative explanations for these inter-relations that future research could identify.

Again, considering the overall correlation results, grammatical competence seems to show the next largest relationship with the measures of coherence, after morphological awareness and vocabulary knowledge. Grammatical competence is the knowledge to produce a well-structured sentence that assists learners to comprehend meaning with ease. Under these assumptions, grammatical competence may be expected to show relationships with coherence as it may help to organise the text. The present as well as past studies such as Ahmad et al. (2019), Garing (2014) and Saeed (2020) found that the organisation of the text bears a high correlation with coherence.

Schleppegrell's (2004) work may help to explain why grammatical knowledge correlates with coherence at a lower level than morphological awareness. She asserted that the grammar of everyday communication differs from that of academic discourse. The latter is more formal, better organised, and more elaborated through nouns and noun phrases rather than clause structures. The development of the academic form of the language may improve cognitive functioning and language/writing practice. However, the English second language participants of this study may not have been proficient enough in the academic form of writing for this to support the development of grammatical competence. Hence, even for those with good coherence scores may not have scored well on the grammar tasks used in the current study, leading to smaller correlations between the writing and grammatical tasks. The prediction here would be that with improved academic skills, more experience of complex sentence structure will follow, which will in turn support grammatical competence. Hence, it is likely that a reciprocal relationship exists such that grammar skills will

support coherent writing, but equally text experience will help improve grammar skills.

However, from the results charted in the present study, phonological and orthographic awareness exhibited a lower size correlation across the four measures of coherence than morphological awareness, vocabulary knowledge and grammatical competence. One possible reason might be related to the methods of teaching English in government institutions in Pakistan (where the research was conducted). Most Pakistani government institutions teach English predominantly through a grammar-translation method (Ahmed, 2019; Shamim, 2008). Given that most English teaching takes place in the local language, learners have less chance of listening to the target language, a factor which may consequently affect their learning of sound patterns in the target language, and their development of phonological awareness in that language. If a skill is not developed, then it would be unlikely to be used in writing production.

The current research also considered the relationships between the linguistic skills measures and the sub-component parts of each measure of coherence. The sub-component parts of Holistic Coherence Scale showed mainly medium size correlations with morphological awareness, vocabulary knowledge and grammatical competence. These were Focus, Organisation and Closure. A good range of vocabulary related to the main topic is required to keep uniformity of the text. It is likely to result in complementary interactions between these skills and Focus by claiming morphological awareness and vocabulary knowledge related to meaning. Organisation of the text refers to the step-by-step description of the events which makes the text logical and meaningful as compared to any random collection of events which makes no sense. Morphological awareness helps to organise the correct word structure by knowing the proper knowledge of affixation and thus enhances vocabulary knowledge. For example, the discourse markers such as 'finally' 'consequently' and some others help to organise the text and need the proper use of affixation.

As a measure of coherence, Topical Structure Analysis exhibited higher relationships with morphological awareness, vocabulary knowledge and grammatical competence than phonological and orthographic awareness. Its sub-components, Sequential progression and Unrelated progression produced larger correlations with morphological awareness and vocabulary knowledge, whereas Parallel progression showed larger correlations with phonological and orthographic awareness. Most of these correlations were small (the largest being .31 between vocabulary and Unrelated progression) with many being insignificant. Based on these findings, it is possible to suggest that the language skills assessed in the current study do not substantially support the production of these types of progressions. The higher association of Sequential progression and Unrelated progression with morphological awareness and vocabulary knowledge indicates that these two progressions are more related to meaning, whereas the higher association of Parallel progression with Cohesion and Organisation, suggests that Parallel progression is more related to the connectivity and structure of the text than to meaning.

Limitations of the study

The present study is limited in terms of using only descriptive essays in the study. Previous studies, such as that of Ghazanfari et al. (2011) have documented the effects of different genre on coherence analysis, pointing out that different genres of writing employ different structure. Therefore, using other genre of essays may produce different results.

The present study administered a general vocabulary measure which showed the highest level of correlation with coherence. Subsequent studies might also meaningfully use depth and width of vocabulary measures separately to investigate as predictors of coherence. It would greatly help in the understanding of coherence as to whether the in-depth understanding of a word is more associated with coherence, or with the quantity of words.

Gender has been documented to have effects on language learning. Female learners have been found to be better language learners than males (Saeed et al., 2011). Future studies could address the effect of gender differences in written coherent text in adult ESL learners' writing and also the reason behind it. For instance, both genders differ in employing strategies to write a coherent text.

CONCLUSION

The purpose of the study was to find out the predictors of coherence in adult ESL learners' writing. The study identified mainly medium-size correlations between the reader-based measures of coherence and the language skills of morphological awareness, vocabulary knowledge and grammatical competence in adult ESL learners' writing. However, the coherence-language skill correlations were generally small in the analyses involving the text-based measures. Phonological and orthographic awareness measures showed mainly small relationships with all coherence measures used in the study. These findings suggest that more meaning-related language skills were more likely to support coherent text production than those focused on linking written and spoken language, though the variability in relationships across the types of coherence measures warrants further research. Some sub-components of the coherence measures did produce medium size relationships with these same meaning-related language skills, but again the main relationships were with sub-components of the reading-based coherence measure. These differences suggest that components of coherence measures may be somewhat independent of those language skills associated with writing ability. Further re-

search identifying the skills that may lead to better scores in coherence assessments, and which support the development of these skills in second language learning contexts, would seem necessary to inform practices used in writing classes.

ACKNOWLEDGEMENT

This study was funded entirely by the researchers themselves.

AUTHOR CONTRIBUTIONS

Abdul Saeed: conceptualization, data curation, investigation, methodology, project administration, validation, visualization, writing-original draft, formal analysis.

John Everatt: conceptualization, formal analysis, methodology, project administration, supervision, validation, visualization, writing-original draft, writing-review & editing.

Amir Sadeghi: conceptualization, methodology, project administration, resources, supervision, validation, writing-review & editing.

Athar Munir: data curation, investigation, methodology, project administration, resources, validation, writing-review & editing.

DECLARATION OF COMPETING INTEREST

None declared.

REFERENCES

Abu-Rabia, S. (2001). Testing the interdependence hypothesis among native adult bilingual Russian-English students. *Journal of Psycholinguistic Research, 30*(4), 437-455. https://doi.org/10.1023/A:1010425825251

Ahmed, F. E. Y. (2019). Errors of unity and coherence in Saudi Arabian EFL university students' written paragraph – A case study of College of Science & Arts, Tanumah, King Khalid University, Kingdom of Saudi Arabia. *European Journal of English Language Teaching*, *4*(3), 125-155. https://dx.doi.org/10.5281/zenodo.321555

Ahmad, M., Mahmood, M. A., & Siddique, A. R. (2019). Organisational skills in academic writing: A study on coherence and cohesion in Pakistani research abstracts. *Languages*, *4*(4), 1-26. https://doi.org/10.3390/languages4040092

Alsagoafi, A. A. (2013). *An investigation into the construct validity of an academic writing test in English with special reference to the academic writing module of the IELTS test.* [Unpublished doctoral thesis]. University of Exeter. https://ore.exeter.ac.uk/repository/handle/10871/10121

Attelisi, A. A. S. (2012). *The impact of teaching topical structure analysis on EFL writing with special reference to undergraduate students in Libya* [Unpublished doctoral thesis]. Newcastle University. https://theses.ncl.ac.uk/jspui/handle/10443/1619

Bamberg, B. (1984). Assessing coherence: A reanalysis of essays written for the National Assessment of Educational Progress, 1969-1979. *Research in the Teaching of English*, *18*(3), 305-319. https://www.jstor.org/stable/40171021

Berninger, V. W., Abbott, R. D., Nagy, W., & Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research, 39*(2), 141-163. https://doi.org/10.1007/s10936-009-9130-6

Berninger, V. W., Nagy, W., & Beers, S. (2011). Child writers' construction and reconstruction of single sentences and construction of multi-sentence texts: Contributions of syntax and transcription to translation. *Reading and Writing, 24*(2), 151-182. https://doi.org/10.1007/s11145-010-9262-y

Candelo, J. E., Soto, J. D., Torres, L., Schettini, N., Calle, M., García, L., & de Castro, A. (2018). Coherence and cohesion issues in argumentation documents written by engineering students. *In proceedings of 2018 IEEE Global Engineering Education Conference (EDUCON)* (pp. 156-160). IEEE. 10.1109/EDUCON.2018.8363222

Chiang, S. (2003). The importance of cohesive conditions to perceptions of writing quality at the early stages of foreign language learning. *System*, *31*(4), 471-484. https://doi.org/10.1016/j.system.2003.02.002

Connor, U., & Lauer, J. (1985). Understanding persuasive essay writing: Linguistic/rhetorical approach. *Text, 5*(4), 309-326. http://hdl.handle.net/1805/2662

Crossley, S.A., & McNamara, D.S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In R. Catrambone & S. Ohlsson (Eds.), *Proceedings of the 32nd annual conference of the Cognitive Science Society* (pp. 984-989). Cognitive Science Society.

de Beaugrande, R. A., & Dressler, W. U. (1981). *Introduction to text linguistics*. Longman.

Diamond, K. E., Gerde, H. K., & Powell, D. R. (2008). Development in early literacy skills during the pre-kindergarten year in Head Start: Relations between growth in children's writing and understanding of letters. *Early Childhood Research Quarterly, 23*(4), 467-478. https://doi.org/10.1016/j.ecresq.2008.05.002

Garing, A. G. (2014, March). Coherence in argumentative essays of first-year College of Liberal Arts students at De La Salle University. In *DLSU Research Congress* (pp. 1-15). DLSU press.

Ghazanfari, M., Alavi, S. Z., & Ghabanchi, Z. (2011). The relationship between types of paragraphs and topic progression used in paragraphs written by Iranian EFL students. *Journal of International Education Research, 7*(4), 39-46. https://doi.org/10.19030/jier.v7i4.8003

Halliday, M. A. K., & Hasan, R. (2014). *Cohesion in English*. Routledge.

Harrison, G. L., Goegan, L. D., Jalbert, R., McManus, K., Sinclair, K., & Spurling, J. (2016). Predictors of spelling and writing skills in first-and second-language learners. *Reading and Writing, 29*(1), 69-89. https://doi.org/10.1007/s11145-015-9580-1

Hewings, A., & North, S. (2006). Emergent disciplinarity: A comparative study of Theme in undergraduate essays in geography and history of science. In Whittaker, R., McCabe, A. & O'Donnell, M. (Eds.). *Language and literacy: Functional approaches* (266-281). Continuum.

Jones, S., Myhill, D., & Bailey, T. (2013). Grammar for writing? An investigation of the effects of contextualised grammar teaching on students' writing. *Reading and Writing, 26*(8), 1241-1263. https://doi.org/10.1007/s11145-012-9416-1

Khalil, A. (1989). A study of cohesion and coherence in Arab EFL college students' writing. *System, 17*(3), 359-371. https://doi.org/10.1016/0346251X(89)90008-0

Kılıç, M., Genç, B., & Bada, E. (2016). Topical structure in argumentative essays of EFL learners and implications for writing classes. *Journal of Language and Linguistic Studies, 12*(2), 107-116. https://dergipark.org.tr/en/pub/jlls/issue/36115/405544

Kim, M., & Crossley, S. A. (2018). Modeling second language writing quality: A structural equation investigation of lexical, syntactic, and cohesive features in source-based and independent writing. *Assessing Writing, 37,* 39-56. https://doi.org/10.1016/j.asw.2018.03.002

Kim, Y.-S., Al Otaiba, S., Sidler, J. F., & Gruelich, L. (2013). Language, literacy, attentional behaviors, and instructional quality predictors of written composition for first grades. *Early Childhood Research Quarterly, 28*(3), 461-469. https://doi.org/10.1016/j.ecresq.2013.01.001

Knoch, U. (2007). 'Little coherence, considerable strain for reader': A comparison between two rating scales for the assessment of coherence. *Assessing writing, 12*(2), 108-128. https://doi.org/10.1016/j.asw.2007.07.002

Lautamatti, L. (1978). Observations on the development of the topic in simplified discourse. *Afinla-import, 8*(22), 71-104.

Lee, I. (2002). Teaching coherence to ESL students: A classroom inquiry. *Journal of Second Language Writing, 11*(2), 135-159. https://doi.org/10.1016/S1060-3743(02)00065-6

Lee, S. H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instrcution. *Syetem, 31*(4), 537-561. https://doi.org/10.1016/j.system.2003.02.004

Mackenzie, N., & Hemmings, B. (2014). Predictors of success with writing in the first year of school. Issues in Educational Research, 24(1), 41–54. https://search.informit.org/doi/10.3316/informit.352676063609817

Masadeh, T. S. (2019). Cohesion and coherence in the writings of Saudi undergraduates majoring in English. *Journal of Social Sciences and Humnaities, 5*(3), 200-208. http://www.aiscience.org/journal/paperInfo/jssh?paperId=4522

McCutchen, D., & Stull, S. (2015). Morphological awareness and children's writing: Accuracy, error, and invention. *Reading and Writing, 28*(2), 271-289. https://doi.org/10.1007/s11145-014-9524-1

McCutchen, D., Stull, S., Herrera, B. L., Lotas, S., & Evans, S. (2014). Putting words to work: Effects of morphological instruction on children's writing. *Journal of Learning Disabilities, 47*(1), 86-97 https://doi.org/10.1177/0022219413509969

McCulley, G. A. (1985). Writing quality, coherence, and cohesion. *Research in the Teaching of English*, *19*(3), 269-282. http://www.jstor.org/stable/40171050

McKenna, M. J. (1988). The development and validation of a model for text coherency. https://eric.ed.gov/?id=ED302830

Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System, 28*(2), 291-304. https://doi.org/10.1016/S0346-251X(00)00013-0

Moore, T. J., & Morton, J. (2007). Authenticity in the IELTS academic module writing test: a comparative study of task 2 items and university assignments. In L. Taylor & P. Falvey (Eds.), *Studies in language testing 19: IELTS collected papers* (pp. 197 - 248). Cambridge University Press.

Müller, A., & Daller, M. (2019). Predicting international students' clinical and academic grades using two language tests (IELTS and C-test): A correlational research study. *Nurse Education Today, 72*, 6-11. https://doi.org/10.1016/j.nedt.2018.10.007

Northey, M., McCutchen, D., & Sanders, E. A. (2016). Contributions of morphological skill to children's essay writing. *Reading and Writing, 29*(1), 47-68. https://doi.org/10.1007/s11145-015-9579-7

O'brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics, 27*(03), 377-402. https://doi.org/10.1017/S0142716406060322

Pearson, W. S. (2018). Written corrective feedback in IELTS writing task 2: Teachers' priorities, practices, and beliefs. *Tesl-Ej, 21*(4), 1-32. https://eric.ed.gov/?id=EJ1172568

RahmtAllah, E. A. E. (2020). EFL students' coherence skill in writing: A case study of third-year students of bachelors in English language. *English Language Teaching*, *13*(8), 120-126. https://doi.org/10.5539/elt.v13n8p120

Roessingh, H., Elgie, S., & Kover, P. (2015). Using lexical profiling tools to investigate children's written vocabulary in grade 3: An exploratory study. *Language Assessment Quarterly, 12*(1), 67-86. https://doi.org/10.1080/15434303.2014.936603

Saeed, A. (2020). Cognitive predictors of coherence in adult ESL learners' writing. (Unpublished doctoral dissertation). University of Canterbury. https://doi.org/http://dx.doi.org/10.26021/2710

Saeed, A., Ghani, M., & Ramzan, M. (2011). Gender Difference and L2 Writing. *International Research Journal of Arts & Humanities (IRJAH), 39*(39). https://sujo-old.usindh.edu.pk/index.php/IRJAH/article/view/1148/1064

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Routledge. https://doi.org/10.4324/9781410610317

Schoepp, K. (2018). Predictive validity of the IELTS in an English as a medium of instruction environment. *Higher Education Quarterly, 72*(4), 271-285. https://doi.org/10.1111/hequ.12163

Shamim, F. (2008). Trends, issues and challenges in English language education in Pakistan. *Asia Pacific Journal of Education, 28*(3), 235-249. https://doi.org/10.1080/02188790802267324

Todd, R. W. (2016). *Discourse topics*. John Benjamins Publishing Company.

van Dijk, T. A. (1977). Semantic macro-structures, knowledge frames, and discourse comprehension. In M. A. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 03-32). Erlbaum.

Wang, Y., Yin, L., & McBride, C. (2015). Unique predictors of early reading and writing: A one-year longitudinal study of Chinese kindergarteners. *Early Childhood Research Quarterly, 32*, 51-59. https://doi.org/10.1016/j.ecresq.2015.02.004

Witte, S. P., & Faigley, L. (1981). Coherence, cohesion, and writing quality. *College Composition and Communication, 32*(2), 189-204. https://doi.org/10.2307/356693

Wong, A. S. C. (2012). *An investigation of the predictors of L2 writing among adult ESL students* [Unpublished doctoral thesis]. University of Canterbury. http://dx.doi.org/10.26021/9887

Zhang, D., & Koda, K. (2012). Contribution of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge and reading comprehension among advanced EFL learners: Testing direct and indirect effects. *Reading and Writing, 25*(5), 1195-1216. https://doi.org/10.1007/s11145-011-9313-z

Zhao, J. (2011). *Spelling English words: Contributions of phonological, morphological and orthographic knowledge in speakers of English and Chinese* [Unpublished doctoral dissertation]. Texas A&M University.

https://doi.org/10.17323/jle.2022.13590

Developing and Piloting a Q-sample on Chinese Language Learners' Epistemic Beliefs

Yanyan Wang [1,2], Larisa Nikitina [1], Jagdish Kaur [1], Fumitaka Furuoka [1]

[1] University of Malaya, Malaysia
[2] Chifeng University, China

ABSTRACT

Background. Epistemic beliefs refer to a person's viewpoints about the nature of knowledge and the process of knowing. A number of studies have explored language learners' subjective views about what knowing and learning a foreign or a second language (L2) means to them personally. For the most part, these studies adopted quantitative research designs and employed self-reported questionnaires with Likert-type scales to collect the data.

Purpose. This pilot study aimed to assess feasibility of adopting Q-methodology (Q) for explorations of language-related epistemic beliefs held by Chinese university students. A detailed account is given of the development of the research instrument (Q-sample); the findings from the Q-sample piloting are reported.

Methods. The newly-developed Q-sample consisting of 45 statements was tested among six students learning the English language in a university in Mainland China. The students were at a different level of the English language proficiency. The 11-point Q-sorting grid had the values ranging from -5 ("Most disagree") to +5 ("Most agree"). To gain deeper insights into the students' personal epistemologies, a semi-structured post Q-sorting interview was conducted with each student.

Results. The newly-developed Q-sample was found suitable for exploring language-related epistemic beliefs. Two groups of students sharing similar beliefs were distinguished. Students who clustered together to form Factor 1 held stronger viewpoints concerning certainty of knowledge; these students were at a lower English language proficiency level. The students who conglomerated on Factor 2 were at a higher level of language proficiency and they held stronger opinions relating to the authority and source of knowledge.

Implications. The findings highlighted the relevance and salience of the epistemic beliefs pertaining to the process of acquiring knowledge. Further research with larger numbers of students is required to explore the role of language proficiency in shaping language learners' personal epistemologies.

KEYWORDS

Q-methodology, Q-sample, epistemic beliefs, English language learners, Chinese students

Citation: Wang Y., Nikitina L., Kaur J., & Furuoka F. (2022). Developing and Piloting a Q-sample on Chinese Language Learners' Epistemic Beliefs. Journal of Language and Education, 8(3), 119-132. https://jle.hse.ru/article/view/13590

Correspondence:
Larisa Nikitina,
larisa@um.edu.my

Received: December 24, 2021
Accepted: September 17, 2022
Published: September 30, 2022

INTRODUCTION

Beliefs that people hold about the nature of knowledge, knowing and the process of acquiring knowledge are referred to as 'epistemic beliefs'. Important epistemological questions concerning language learning and knowing have been raised by theoretical linguists. Noam Chomsky, for example, sought the answers to the questions such as "What constitutes knowledge of languages?" and "How is knowledge of language acquired?" (Chomsky, 1986, p. 3). As Chomsky proposed, if a person knows a language he or she "has mastered a set of rules and principles that determine an infinite, discrete set of sentences, each of which has a fixed form and a fixed meaning or meaning potential" (Chomsky, 1975, p.303). He referred to this "set of rules and principles" that govern the formation of a meaning as the generative grammar. However, at a layman's level, each person who learns – or has learnt – a new language would have his or her

own conceptions and beliefs pertaining to these deep philosophical questions. These beliefs can be approached and studied as "socially-shared intuitions" (Jehng, Johnson & Anderson, 1993, p. 24) about the nature of knowing and learning a foreign or a second language (L2). There is only a handful of studies that explored epistemic beliefs held by language learners. In most of these studies, researchers conducted questionnaire surveys to collect the data (e.g., Kahsay, 2019; Mori, 1999; Nikitina & Furuoka, 2018). Only in rare instances other methodologies, such as Q-methodology, were employed (e.g., Rock, 2013).

This study aimed to assess feasibility of adopting Q-methodology (Q) for explorations of a complex concept of epistemic beliefs among university students learning a foreign (English) language. To achieve this aim, and considering that empirical studies on language-related epistemic beliefs are scarce, this article provides a detailed account of developed an instrument (Q-sample) on epistemic beliefs related to learning an additional language. It then proceeds to report the findings from a pilot study that tested the newly-developed Q-sample with a group of English language learners in a university in Mainland China.

BACKGROUND

Q methodology in L2 research on beliefs

Beliefs is an elusive psychological construct; it has been defined in various ways. An influential study by Pajares (1992) proposed that beliefs "travel in disguise and often under alias – attitudes, values, judgements, axioms, opinions, ideology, perceptions, conceptions, conceptual systems, preconceptions, dispositions, implicit theories, personal theories, internal mental processes, action strategies, rules of practice, practical principles, perspectives, repertoires of understanding, and social strategy, to name but a few" (p. 309).

Literature on beliefs in the fields of general education and L2 research is extensive. Various methodological approaches have been adopted to study this topic but quantitative studies that derived their data from Likert-type questionnaires remain prevalent (Barcelos & Kalaja, 2011). Considering the construct's complexity, its rootedness in the context and inherent subjectivity, Q-methodology would offer researchers and language educators particularly rich affordances in their explorations of language learners' – and their teachers'– beliefs. However, the method was rarely employed until very recently. Among the latest studies, Q was adopted to explore language educators' and pre-service teachers' beliefs about multilingualism (Lundberg, 2019, 2020), their pedagogical practice (Lu, Zou & Tao, 2020) and teaching competencies (Irie, Ryan & Mercer, 2018). We were able to locate only one Q study done by Rock (2013) that focused specifically on *epistemic* beliefs. This gap in research literature could be due to a complexity of identifying appropriate areas and dimensions within language-related epistemic beliefs.

In general education research, personal epistemologies and epistemic beliefs have been viewed as consisting of discrete but interrelated dimensions. This was initially proposed by Perry (1970) in a series of his pioneering studies. Hofer (2000) conceptualized personal epistemologies as consisting of two vast areas which pertain to "the *nature of knowledge*" (i.e., what individual people believe constitutes knowledge) and "the nature or *process of knowing*" (i.e., how we come to know what we know) (p. 380, emphasis in original). Furthermore, the 'nature of knowledge' dimension comprises beliefs about certainty and simplicity of knowledge, while the 'process of knowing' dimension relates to the sources and justification of knowledge. At the same time, as it is noted in the research literature (Hofer, 2000; Stahl & Bromme, 2007), beliefs related to the '*process of knowing*' (i.e., the gaining of knowledge or learning) have been viewed as peripheral and they were often excluded by researchers from their studies. To dispute this opinion, Nikitina and Furuoka (2018) referenced Pritchard's (2006) arguments on the importance of procedural (i.e., "know-how" knowledge). Pritchard maintained that "to have knowledge, one's success must genuinely be the result of one's efforts" (p.6) to seek and gain knowledge. Considering that the empirical findings of earlier studies (Mori, 1999; Rock, 2013) as well as their own research on L2 learners' personal epistemologies supported the legitimacy of beliefs about learning (i.e., the 'know-how' beliefs), Nikitina and Furuoka called for a reconceptualization of epistemic beliefs in L2 research. Acknowledging these arguments, the conceptual framework that guided our efforts to develop the concourse and Q-sample on L2 learners' epistemic beliefs includes the 'process of learning' dimension.

Q-Methodology: Its Roots and Main Features

Q-methodology (Q), an approach to investigating individual people's viewpoints and subjectivity on any topic, phenomenon or event, was invented in the 1930s by British physicist and psychologist William Stephenson (1902–1989). Stephenson (1935a) introduced Q in a brief note published in the journal *Nature*. More extensive explanations of the method can be found in Stephenson's numerous writings, including his seminal book *The study of behavior; Q-technique and its methodology* published in 1953. Very briefly, Stephenson noted a contradiction in contemporary psychology research on individual differences where the methods employed by researchers, such as factor analysis of data collected through Likert-scale type surveys (R-methodology), actually yielded the findings on the latent structure of a study's *variables* for an average person in a bigger population – and not the insights on an *individual person* or groups of people with their unique individual conceptions and opinions. Stephenson (1953) proposed to 'reformulate' the

factor analysis so that people – and not the variables – are grouped together based on similarities and differences of the individual viewpoints (Stephenson, 1935b, 1936, 1953). With the main focus on unique and subjective, Q does not require many participants. There are single-case Q studies where several Q-sorts were collected from the same person (see Stephenson, 1953, 1992).

Over decades, Q methodology has gained recognition and popularity in a wide range of academic disciplines, including education research, environmental sciences, nursing, economics, political science and management. Some Q studies have been done in applied linguistics and L2 research (Caruso & Fraschini, 2021; Damio, 2018; Damio & Hashim, 2014; Fong, 2021; Irie, 2014; Irie & Ryan, 2015; Rock, 2013; Zheng, Lu & Ren, 2019). Several detailed guides on how to conduct a Q study are available and they list typical stages in a Q study (Brown, 1980, 1993; McKeown & Thomas, 2013; Stephenson, 1953; Watts & Stenner, 2012). Generally, upon deciding a topic or phenomenon of interest a researcher would (1) compile a concourse (or wide array of 'communications' on the topic of interest), (2) select a smaller Q-sample from the concourse, (3) identify appropriate participants (P-set) in their study, (4) ask the participants to sort the Q-sample items (to get the Q-sorts that present subjective viewpoints of the people), (5) perform the factor analysis of the Q-sorts, (6) interpret the results of the statistical analysis and analyze other data (e.g., the respondents' further clarifications during interviews or their answers to written open-ended questions) and (7) produce a narrative account.

While designing our study, a search of literature revealed that there is a lack of information on developing a Q-sample on epistemic beliefs in L2 research. It should be noted that one of the main points raised by critics of Q is a perceived lack of methodological rigour and transparency in constructing a systematic Q-sample. Watts and Stenner (2012) referenced Block (2008) who opined that "a set of Q items typically is quickly assembled, structured a priori (often questionably) by the investigator, and is not itself further evaluated as to its sufficiency of meaning" (p. 110). To take note of this issue, this article provides a detailed explanation of the steps we followed and the decisions we made while compiling the concourse and selecting the ensuing Q-sample on language learners' epistemic beliefs. The research question that necessitated the development of the Q-sample was "What epistemic beliefs do Chinese language learners hold about knowing and learning the English language?" To the best of our knowledge, this is the first study that explicitly explains intellectual and decision-making processes while constructing a Q-sample on language learners' epistemic beliefs. The following sub-sections give a detailed description of the concourse development.

METHODS

Participants and Setting

This pilot study was conducted with six students learning English in a large public university in Mainland China, where the first author teaches. A Q study does not require a large number of participants; some studies had as little as six, seven and nine Q-sorters (see Lundberg, 2020). Therefore, six students was a sufficient number of Q-sorters for the purpose of piloting the research instrument.

The six students volunteered to take part in the pilot study by responding to an invitation given to all students in the first authors' English language class. The six participants were very similar to the intended participants in the main study: their age was between 19 and 20 years old and they were majoring in engineering. Even more importantly for a Q study (see Watts & Stenner, 2012), the participants were competent and well-positioned to express their viewpoints on language-related epistemologies as they all had an extensive experience of learning English.

Among the volunteer Q-sorters, three students had a higher level of the English language proficiency and other three students had a lower proficiency level. This helped to achieve some diversity of the P-set. The level of language proficiency was measured by the students' average scores on the two latest English language tests they had taken. The following subsections offer a detailed description of the steps and stages in the development of the Q-sample.

Compiling the Concourse on Language Learners' Epistemic Beliefs: Theoretical and Practical Considerations

A Q study begins with compiling a concourse, which is a wide-ranging "universe of statements for any situation or context" (Stephenson, 1986a, p. 37). Stephenson viewed the process of a concourse development through the prism of a "working theory" of communication where innumerable number of "messages" or communications about a topic are assembled in a concourse along "some broad lines" (p.43). Reminding that the word originates from the Latin *concursus* which means "running together" – of ideas and subjective opinions in a Q study – Stephenson emphasized not only the informational but also the conversational functions and qualities of a concourse. In short, a concourse need to combine a wide range of ideas that "run together" on the topic of interest. Thereafter, a Q-sample consisting of a smaller number of items is drawn from the concourse. Notably, besides textual statements a concourse may consist of drawings, photographs, paintings, music pieces and even scents and fragrances.

As a 'universe' of communications on a topic, a concourse has no limits for the number of statements it can comfortably accommodate. However, all of the statements must be referrals to the object of a study and thus clearly belong to this particular 'universe'; moreover, they must enable the respondents to share their own feelings about the object or phenomenon under study and provide ample affordances for the respondents' expressing their subjective viewpoints (Brown, 1980; Watts & Stenner, 2012). In short, while developing a concourse researchers must ensure that they preserve a 'self-reference' quality of the concourse, and of the Q-sample that stems from it. As Stephenson (1988/89) put it, a concourse "is a hot bed of self-referential potentials" (pp. 7-8 as cited in Watts & Stenner, 2012, p.34).

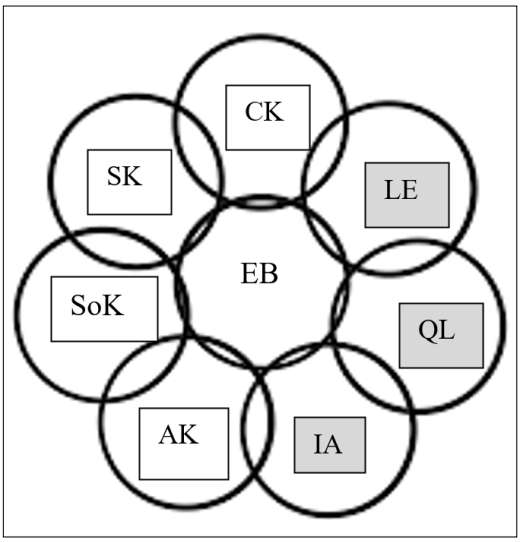
Stephenson (1986) stated that "a concourse is arrived at empirically" (p. 44). Statements in a concourse are sourced from a variety of sources, including scholarly literature (e.g., books, articles and ready-made questionnaires), official documents, the mass media, the social media, interviews and focus groups with experts and potential participants in a Q study and even from every-day conversations and discourses (Brown, 1993; McKeown & Thomas, 2013). In addition, developing a concourse requires including conversational elements that would enable the individual people to express their viewpoints. Researchers may want to adopt a conceptual or a theoretical framework to guide the development of the concourse (Watts & Stenner, 2012). In this approach, dimensions or sub-themes in the matter of a study's interest are firstly identified. Having a framework could enhance credibility of the process of item selection and also lessen the possibility of omitting important aspects concerning the study's subject.

Due to a complexity of the topic 'epistemic beliefs', we decided to have a conceptual framework to guide our concourse building efforts (see Figure 1). For this, important literature on epistemic beliefs in the field of education was consulted, including studies that conducted surveys using self-reported Likert-type scales (Chan & Elliott, 2002; Hofer & Pintrich, 1997; Jehng et al., 1993; Schommer, 1998; Schraw, Bendixen & Dunkle, 2002; Wood & Kardash, 2002). We also consulted very scarce studies on language learners' epistemic beliefs (Mori, 1999; Nikitina & Furuoka, 2018; Rock, 2013).

Based on the literature review presented earlier in this article, the framework to guide the selection of items for the Q-sample incorporated three main dimensions relating to epistemic beliefs, namely, (1) the nature of knowledge, which pertains to beliefs in simplicity and/or certainty of knowledge and knowing, (2) the nature of knowing, which concerns beliefs regarding the source of knowledge and a belief in an omnipresent authority of knowledge and (3) the nature and process of learning, which incorporates beliefs pertaining to quick learning, learning effort and a belief in innate or fixed ability to learn, as shown in Figure 1.

Figure 1

Conceptual framework for the concourse development



Notes: EB = Epistemic Beliefs; CK = Certainty of Knowledge; SK = Simplicity of Knowledge; SoK = Source of Knowledge; AK = Authority of Knowledge; IA = Innate/Fixed Ability; QL = Quick Learning; LE = Learning Effort. The shaded boxes indicate the learning-related dimensions.

In total, 290 statements on epistemic beliefs were sourced from the empirical studies. In addition, in order to include the "conversational" element in the concourse, we collected the opinions of English language learners and teachers in China. The students were asked the open-ended question "What does knowing a foreign language mean personally to you?". The discussions with the language instructors centered on the question "What do we know when we know a foreign language (English)?". This step yielded another 95 items to the concourse. Altogether, we gathered 385 statements pertaining to epistemic beliefs in the concourse (see Table 1).

Upon further scrutiny, we noted that not all of the statements provided by the students pertained to personal epistemologies. Furthermore, there were statements that did not align with the conceptual framework depicted in Figure 1. For example, there were several statements that concerned the importance of learning the target language culture. These statements were removed. In addition, we noted that the statements collected from the various questionnaires were often recurring, i.e., a different wording was used to express essentially the same point. As advised in methodological literature on Q (Watts & Stenner, 2012), the repetitive items were discarded. As a result, there remained 200 workable statements in the concourse pool where 157 statements came from the questionnaires and 43 statements were obtained during the brainstorming sessions with the language learners and discussions with English language instructors in China. Each of the remaining statements aligned well with one of the dimensions in the conceptual framework presented in Figure 1.

The concourse serves as a raw material in a Q study. It usually contains a large number of statements, which would be too cumbersome, if not impossible, for the respondents (P-set) to sort. Therefore, a representative subset of 40 to 80 statements called 'Q-sample' needs to be drawn from the concourse. As Watts and Stenner (2012) noted, the Q-sample "need only contain a representative condensation of information to do its job effectively" (p. 65). The next subsections give a detailed information on the Q-sample development.

Constructing the Q-Sample

To yield meaningful insights, a much condensed Q-sample must preserve the breadth and depth of the original concourse (Brown, 1980; Stephenson, 1953; Watts & Stenner, 2012). Methodological literature distinguishes two types of Q-samples – structured and unstructured (Watts & Stenner, 2012). The structured approach requires that the Q-sample statements cluster around pre-determined theoretically based dimensions, which was the case in the current study. An approximately equal number of items should capture each dimension. The main advantage of adopting a structured approach is that it helps ensuring that the Q-sample is balanced and representative (Watts & Stenner, 2012). The unstructured approach does not require that the components or sub-themes are predetermined.

The number of statements in the Q-sample needs to be manageable and usually the number would be between 40 and 80 items, though it is possible to have a smaller Q-sample (Brown, 1980; Watts & Stenner, 2012). In order to retain a good balance and clarity of the finalized Q-sample, we removed the statements that could be difficult or unclear for the sorters as well as the repetitive or very similar in essence statements. We took particular care to retain statements that are likely to draw emotional reaction from the participants and that retained a 'conversational' element (Stephenson, 1953; Watts & Stenner, 2012). As a result, the concourse of 200 workable statements was reduced to 54 statements. At this stage, we refined the wording of the statements to reflect the object of the study, namely, epistemic beliefs held by language learners. A particular care was taken to ensure that (1) each remaining item fits one of the epistemic beliefs dimensions depicted in Figure 1, (2) each dimension contains approximately the same number

Table 1

Sources for concourse development

Types of information	Sources / persons	Number of statements
Questionnaires on general epistemic beliefs	Schommer (1998)	63
	Schraw et al. (2002)	32
	Wood and Kardash (2002)	38
	Jehng et al. (1993)	15
Questionnaires on language learners' epistemic beliefs	Mori (1999)	53
	Rock (2013)	36
	Nikitina and Furuoka (2018)	53
Brainstorming sessions with students in China	10 volunteer students	89
Personal discussions with language instructors in China	2 language instructors	6
Total		**385**

Table 2

Q-sample statements and their domain mapping

Dimensions	Sub-dimensions	Number of statements
Nature of Knowledge	Simple knowledge/Seek single answer	7
(*n*=15)	Certain knowledge	8
Nature of Knowing	Authority to knowledge	9
(*n*=15)	Sources of knowledge	6
Learning Process	Innate ability	4
(*n*=15)	Quick learning	7
	Learning effort	4
Total		**45**

of items and (3) the statements are relevant in a Chinese educational context, the Q-sample was further reduced to include 45 statements (see Table 2). An English version of the Q-sample is included in the appendix.

As can be seen from the table, there was some variability in the number of statements in the sub-dimensions. For example, among the 15 statements in the Learning Process dimension, 4 statements related to innate ability for learning a foreign language, 7 statements pertained to the belief in quick learning while 4 statements concerned the learning effort. Such mild imbalances are acceptable since, overall, the Q-sample was representative of the concourse and it retained the theoretical depth and breadth (Stephenson, 1953; Watts & Stenner, 2012).

When the Q-sample was finalized, the statements were translated from English to Chinese. We followed recommendations in the methodological literature (Dörnyei & Taguchi, 2009) to ensure that the statements in the translated version sound natural in Chinese and that no statements would invite idiosyncratic interpretations by the respondents. Two language experts were invited to verify the Chinese version.

Evaluating and Refining the Q-Sample

Two domain experts were invited to evaluate the Q-sample. Based on their feedback, 2 statements were deleted from the proposed Q-sample. These were the statements "Knowledge of English is tentative and evolving rather than static" and "Native speakers of English are the source to seek help when students encounter difficult problems with English". Two new statements were added, namely, "The best thing about English knowledge is that there is always one right answer" and "Language learners who disagree with native speakers about grammar or vocabulary usage are over-confident". The wording of 18 statements was refined. For example, the item "Memorizing vocabulary and grammar is all that is needed to perform well in exams" was reworded as "Memorizing vocabulary and grammar is all that is needed to know English". This re-oriented the statement's focus toward epistemic matters from its initial concerns with their exam performance. Accordingly, some changes were done to the translated version of the Q-sample (see Table 3). Following this, a language expert was invited to do reverse translation.

We then prepared 45 cards with the statements and proceeded to piloting the Q-sample. The next section gives more details of the pilot study.

Piloting the Q-Sample

The six students who had volunteered to take part in this pilot study were instructed to place each card in the distribution grid prepared by the first author. The prompt given to the Q-sorters was as follows: "People have various ideas and opinions of what it means to know a foreign language, such as the English language. These cards contain some of these opinions. With which opinions do you agree? With which do you disagree? About which statements do you feel less certain? Please sort the cards accordingly". As noted in methodological literature on Q, the individually held viewpoints are expressed and collected during the process of Q-sorting.

We used 11-point sorting grid for the 45-item Q-sample. The values ranged from -5 ("Most disagree") to +5 ("Most agree"). Regarding the grid for Q-sorting, researchers would need to decide whether its shape (i.e., the curve or slope) should be steeper or flatter. Brown (1980) recommended a steeper distribution if the topic is not very familiar to the participants; this way more of the Q-sample items could be placed at or near the middle of the grid. In our study, we targeted the participants with an extensive experience of learning a foreign (English) language. Such participants would have well-formed subjective views concerning knowing – and learning – a new language. Therefore, we decided to have a flatter (*platykurtic*) distribution of the Q-sort grid; this also would allow for a more granular view of the respondents' 'extreme' opinions (i.e., the statements placed at or near the end points of the grid) (Brown 1980; Watts

Table 3

Examples of revisions in the English–Chinese translation

English statements	Original Chinese statements	Revised Chinese statements
English language never changes. In the future, it will be the same as today.	英语永远不会改变：现在如此，将来也会如此。	英语语言永远不会改变：将来的英语与现在的英语会完全相同。
I never doubt information about English language use that I receive from native speakers of English.	我从不怀疑从以英语为母语的人那里获得的有关英语使用的信息。	我从不怀疑从英语本族人那里获得的有关英语使用的信息。
To gain knowledge you need to discover how to learn.	要获得知识，你需要发现如何学习。	要获得知识，你需要找到如何学习的方法。
People can study English language for years and still not have a good knowledge of it.	人们可以学习英语多年，但仍然不了解这门语言。	有人可能学习英语多年，但仍然不能把这门语言学好。

& Stenner, 2012). The optimal slope (*kurtosis*) of the distribution grid (see Figure 2) was calculated using EViews 10.0 statistical package.

As Brown (2008) noted, "There is no correct way to do a Q sort" (p. 700) because a completed Q-sort would reflect each individual participants' unique subjective opinion. However, in order to give some 'technical' guidance on the sorting process, we designed a booklet with instructions to the participants. We also prepared open-ended questions for the post-Q sorting interviews (e.g., "Could you explain why you have most agreed with these statements?", "What about these statements?", "Would you like to add anything else?").

It took between 30 and 45 minutes for each participant to complete the Q-sort. The researcher was present at all times during the sorting in order to be available to answer any possible queries from the participants. Each completed Q-sort configuration was photographed to retain it for further analysis. Following this, a brief interview was conducted during which the participants were asked to share their general views on the study topic and explain the configurations they had produced, particularly pertaining to the items placed at or just next to the extreme ends of the grid. Each interview lasted about 20 minutes.

After all the Q sorts had been collected, the numerical data were keyed-in into a Microsoft Excel file to enable further statistical analysis where the Q sorts were firstly inter-correlated and then subjected to a factor analysis which involved factor extraction and factor rotation. The factor analysis in Q places respondents with similar views (as evidenced by their Q sorts) in the same factor (Brown 1980; Stephenson 1953; Watts & Stenner, 2012). According to Stephenson (1986a), the purpose of the factor analysis in Q is to distinguish "schemata" or an individual's subjective viewpoints concerning the study's matter. Hence, Q factors "are indicative of schemata" (p. 53) which are derived from a Q-sample by each individual Q-sorter. These schemata allow researchers to penetrate further into each individual respondent's subjectivity because the schemata are developed from the individual person's past experience and present situation (see Stephenson, 1986a for further discussion). Since all schemata are highly subjective, the researcher would need to use abductive reasoning and try to situate oneself in the mind of the Q-sorters when interpreting a Q factor (p. 54).

Besides, the researcher would need to make decisions regarding the statistical procedure and software. After the correlation analysis of the Q-sorts was done, this study performed centroid factor analysis (Brown, 1980) with varimax rotation. KADE software (Banasick, 2019) designed specifically for analysing Q-sorts aided in the construction of the factor arrays. A factor array is a single Q sort that represents the essence of a relevant factor; one factor array is produced for each of the factors that transpired in the course of the analysis. A Python code was written to visualize the findings from the factor rotation with regard to the students' academic achievement.[1] The findings from the pilot study are reported in the next section.

RESULTS

The results from the correlation analysis are shown in Table 4. Higher correlations indicate a higher degree of similarity in the respondents' subjective opinion, such as their epistemic beliefs. As can be seen from the table, the correlations among the Q-sorts were rather high, except for Participant 3 (P3). Table 5 shows the findings on the factors extracted by the centroid method. All of the six participants' Q-sorts had high loadings on Factor 1. This factor has an eigenvalue of 3.267 and explains 54% of the total variance. Only Q-sorts obtained from Participant 1, Participant 4 and Participant 5 had positive loadings on Factor 2 that has an eigenvalue 0.292 and explains 5% of the study variance. These results suggest the existence of two groups of language learners whose epistemic beliefs are similar.

Table 6 shows rotated with the varimax procedure factors and their loading. The varimax procedure suggested a 45 degree clock-wise rotation. As Table 6 shows, each factors had Q-sorts with significant loading, which was calculated as 0.39 in this study (the formulas for this calculation can be found in Brown, 1980, pp. 222-223). However, there were two confounded Q sorts (1 and 6) and one non-significant Q-sort (4).

The findings concerning the factors' eigenvalues (which are recommended to exceed 1.00) and their significant loadings (at least two such loadings on a factor are desirable) can help researchers to make decisions about the number of factors to retain for a further analysis. The final decision, however, rests with the researcher (Watts & Stenner, 2012). This is because this decision would be rooted in the researcher's own knowledge, his or her understanding of the situation and participants and it will be guided by abductive reasoning with regard to the data. We decided to retain two

Figure 2

Q-sort grid distribution

Value	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5
No. of statements	2	3	4	5	5	7	5	5	4	3	2

[1] The Python code is available at: https://sites.google.com/view/homepageforeconometricpython/paper-1

factors for further analysis. The two factors accounted for 59% of total variance. We also took note that the presence of the confounded and non-significant Q-sort loadings indicated the need for improvements in the Q-sample context. To gain further insights, factor arrays for each of the two factors were produced, as shown in Figure 3 and Figure 4. -

The findings revealed that the students' personal epistemologies were rather uniform: the majority of the statements in the two factors were consensus statements. These statements are highlighted in colour. For example, the opinions that knowledge of English cannot be acquired at once (statement #39) received a strong endorsement. Quite logically, there was a disagreement with the opinion that is contrary to this viewpoint (statement #38).

Furthermore, both factors expressed a strong support of the opinion that even if someone has a good knowledge of English there would remain many things to be learned (statement #10), that one needs to discover how to learn (statement #30) and that one needs to persevere in learning English (statement #44). At the same time, the two factors conveyed a strong disagreement that talent plays the major role in learning English (statement #43).

Next, in order to identify the main characteristics of each factor we compared their distinguishing statements, as recommended in methodological literature (Albright, Christofferson, McCabe & Montgomery, 2019; Watts & Stenner, 2012). This analysis indicated that Factor 1 affirmed a collective belief in certainty of linguistic knowledge. The statement that linguistic knowledge is certain as it has been agreed upon by linguists and language experts was a distinguishing statement which received a comparatively strong endorsement in Factor 1 (15; +2). Accordingly, a statement contrary to this viewpoint was among the distinguishing statements that had a strong disagreement (8;-3). As Participant 2, whose Q-sort was strongly associated with Factor 1 and thus represents an exemplar opinion of this group of students, explained,

> *"As I see it, a knowledge of English language that language learners receive must be established knowledge and not uncertain one."*

Factor 2 reflects a greater endorsement of the opinions regarding the source and authority of knowledge. Its distinct viewpoint was that the source of linguistic knowledge mainly comes from the language teachers (27; +2). Interestingly, with regard to the authority of knowledge, the students who conglomerated on this factor found acceptable to doubt

Table 4

Findings from the correlation analysis

	P1	P2	P3	P4	P5	P6
P1	1.00	0.62	0.52	0.65	0.69	0.63
P2	0.62	1.00	0.48	0.53	0.48	0.64
P3	0.52	0.48	1.00	0.36	0.46	0.50
P4	0.65	0.53	0.36	1.00	0.61	0.41
P5	0.69	0.48	0.46	0.61	1.00	0.50
P6	0.63	0.64	0.50	0.41	0.50	1.00

Note: P = participant / sorter

Table 5

Findings from centroid factor analysis

Sorters	Factor 1	Factor 2
P1	0.873	0.132
P2	0.752	-0.212
P3	0.614	-0.131
P4	0.654	0.273
P5	0.744	0.238
P6	0.728	-0.282
Eigenvalue	3.267	0.292
Variance explained	54%	5%

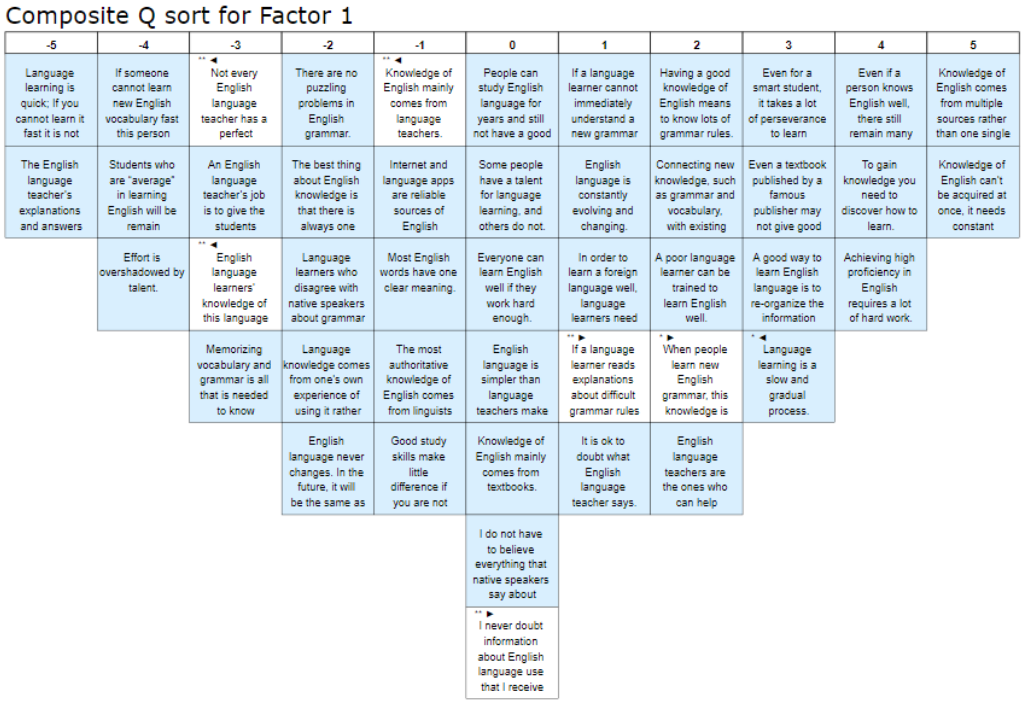
Table 6

Rotated factors and their loadings

Sorters	Factor 1	Factor 2
P1	0.533	0.704
P2	0.687	0.371
P3	0.531	0.334
P4	0.301	0.273
P5	0.367	0.649
P6	0.718	0.690

Figure 3

Factor array for Factor 1



Note: All Q-sample statements were in Chinese. Consensus statements are highlighted in colour.

Factor 1

Threshold	Z-score	Q Sort Value	Number	Statement
P < 0.005	-1.18	-3	8	English language learners' knowledge of this language can never be certain.
P < 0.05	0.815	2	15	When people learn new English grammar, this knowledge is certain and has been agreed upon by linguists and language experts.
P < 0.01	-0.267	0	17	I never doubt information about English language use that I receive from native speakers of English.
P < 0.0005	-1.019	-3	19	Not every English language teacher has a perfect knowledge of the language they teach.
P < 0.0005	-0.297	-1	27	Knowledge of English mainly comes from language teachers.
P < 0.05	0.969	3	36	Language learning is a slow and gradual process.
P < 0.01	0.496	1	40	If a language learner reads explanations about difficult grammar rules many times, he or she will be able to understand and learn these rules well.

information about the use of English received from native speakers of this language (17; -2). As respondent 1, who was strongly associated with this factor, elucidated,

> *"In contrast to the older days, Chinese language learners today do not blindly believe in so-called 'authority' of language regardless whether it is a native or non-native speaker of English. Everyone makes mistakes no matter who they are".*

Next, we considered the findings regarding similarities and differences in the epistemic beliefs between language learners with a different level of academic achievement; the findings are visualized in Figure 5. As the figure shows, the Q-sorts obtained from the language learners whose academic performance was lower tended to have higher loadings on the first factor; only one such Q-sort had a low loading of 0.3. At the same time, these students' Q-sorts tended to have relatively low factor loadings on the second factor, except for one Q-sort with the 0.65 factor loading.

In short, the factor analysis indicated the language learners tended to share the epistemic beliefs with their peers at the same level of academic achievement. Overall, the findings supported suitability of the Q-sample for gathering subjective views and epistemic beliefs concerning knowing and learning a foreign language among Chinese language learners. We also received some valuable feedback from the participants. For example, several students asked to explain the statement "A good way to learn English language is to re-organize the information according to one's own personal understanding". Also, the students commented that some statements were repetitive. Based on these remarks, some modifications to the Q-sample were done in the following main study.

As to the post Q-sorting interview protocol, though the respondents had no difficulty to answer the questions we still needed to develop a more focused interview and include probing questions of various types. More importantly, when

Figure 4

Factor array for Factor 2

Composite Q sort for Factor 2

-5	-4	-3	-2	-1	0	1	2	3	4	5
Effort is overshadowed by talent.	English language never changes. In the future, it will be the same as	There are no puzzling problems in English grammar.	Language knowledge comes from one's own experience of using it rather	The most authoritative knowledge of English comes from linguists	Knowledge of English mainly comes from textbooks.	Having a good knowledge of English means to know lots of grammar rules.	English language teachers are the ones who can help	Even for a smart student, it takes a lot of perseverance to learn	Even if a person knows English well, there still remain many	Knowledge of English can't be acquired at once, it needs constant
The English language teacher's explanations and answers	Language learning is quick; If you cannot learn it fast it is not	Language learners who disagree with native speakers about grammar	Good study skills make little difference if you are not	** ◀ If a language learner reads explanations about difficult grammar rules	If a language learner cannot immediately understand a new grammar	Everyone can learn English well if they work hard enough.	In order to learn a foreign language well, language learners need	To gain knowledge you need to discover how to learn.	Knowledge of English comes from multiple sources rather than one single	* ▶ Language learning is a slow and gradual process.
	If someone cannot learn new English vocabulary fast this person	An English language teacher's job is to give the students	The best thing about English knowledge is that there is always one	I do not have to believe everything that native speakers say about	Some people have a talent for language learning, and others do not.	Connecting new knowledge, such as grammar and vocabulary, with existing	** ▶ Knowledge of English mainly comes from language teachers.	English language is constantly evolving and changing.	Achieving high proficiency in English requires a lot of hard work.	
		Students who are "average" in learning English will be remain	Memorizing vocabulary and grammar is all that is needed to know	** ▶ English language learners' knowledge of this language	** ◀ When people learn new English grammar, this knowledge is	A poor language learner can be trained to learn English well.	Even a textbook published by a famous publisher may not give good	A good way to learn English language is to re-organize the information		
			** ◀ I never doubt information about English language use that I receive	Internet and language apps are reliable sources of English	** ▶ Not every English language teacher has a perfect	People can study English language for years and still not have a good	It is ok to doubt what English language teacher says.			
					Most English words have one clear meaning.					
					English language is simpler than language teachers make					

Note: All Q-sample statements were in Chinese. Consensus statements are highlighted in colour.

Factor 2

Threshold	Z-score	Q Sort Value	Number	Statement
P < 0.005	-0.338	-1	8	English language learners' knowledge of this language can never be certain.
P < 0.05	0.094	0	15	When people learn new English grammar, this knowledge is certain and has been agreed upon by linguists and language experts.
P < 0.01	-1.011	-2	17	I never doubt information about English language use that I receive from native speakers of English.
P < 0.0005	0.033	0	19	Not every English language teacher has a perfect knowledge of the language they teach.
P < 0.0005	0.695	2	27	Knowledge of English mainly comes from language teachers.
P < 0.05	1.652	5	36	Language learning is a slow and gradual process.
P < 0.01	-0.286	-1	40	If a language learner reads explanations about difficult grammar rules many times, he or she will be able to understand and learn these rules well.

keying-in the Q-sort data, we discovered that some photos of the completed Q sorts were not very sharp and it was difficult – though still possible – to read the statements. Originally, we had decided against putting numbers at the back of the cards as we felt this might misguide the sorters. However, after encountering this problem we realized that the numbers at the back of the cards were very much needed to avoid jeopardizing the keying-in of the data in the main study. Otherwise, no major issues were encountered during the study.

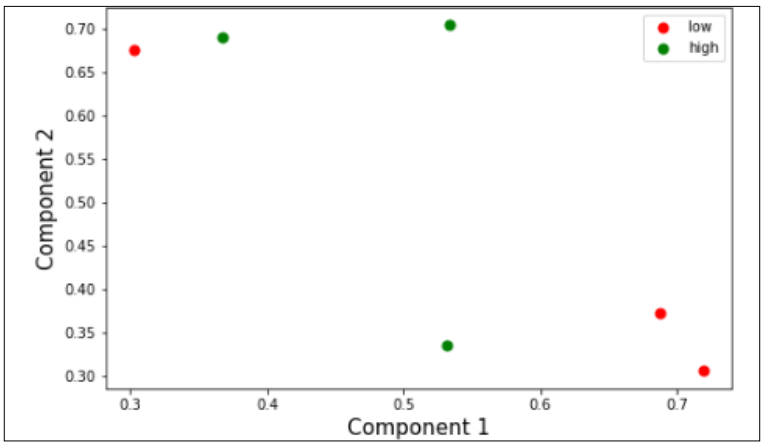
DISCUSSION

Studies on epistemic beliefs are abundant in education research. For the most part, these studies adopted quantitative research approaches (Chan & Elliott, 2002; Hofer & Pintrich, 1997; Jehng et al., 1993; Schommer, 1998; Schraw, Bendixen & Dunkle, 2002; Wood & Kardash, 2002). Q methodology has been rarely adopted in explorations of students' personal epistemologies. In L2 research literature, only one study by Rock (2013) investigated language learners' vocabulary-related epistemic beliefs. The Q study reported in this article is broader in scope as it assessed a variety of language-related epistemic beliefs that people learning a new language might hold. Its findings suggest the existence of two dimensions within the language learners' epistemic beliefs. Students clustering on Factor 1 gave a greater endorsement to the views concerning certainty of knowledge, while those conglomerating on Factor 2 held distinct beliefs relating to the authority and source of knowledge.

The findings also indicated that the epistemic beliefs did not diverge greatly at the intra-individual level as the majority of the statements forming each factor were consensus statements. Notably, the viewpoint that learning a new language, such as English, is a slow and gradual process received a strong endorsement from the students. This is an important

Figure 5

Visualisation of factor loadings



Note: the rotated factor loadings are shown

result stemming from this pilot study. It highlights the relevance and salience of the epistemic beliefs concerning the procedural knowledge and gives support to the arguments advanced in the earlier theoretical and empirical studies on the need to consider such beliefs as integral part of individual people's personal epistemologies (Mori, 1999; Nikitina & Furuoka, 2018; Pritchard, 2006).

As to the role of the language learners' proficiency in English, the findings indicated that the students at the same proficiency level tended to share similar epistemic beliefs. The Q-sorters at a lower level of the language proficiency were associated with Factor 1, while their more proficient peers shared the viewpoints expressed by Factor 2. However, due to a small number of the participants and a lack of prior studies on this phenomenon the conclusions regarding the role of language proficiency in shaping the epistemic beliefs structure must be taken with caution. Further research would be required to gain deeper insights.

A considerable attention has been given in this article to developing the research instrument (Q-sample). As noted in the methodological literature on Q, compiling a concourse and constructing a Q-sample requires perseverance and effort (Brown, 1980; Watts & Stenner, 2012). In the case of the current study, the Q-sample development had lasted 6 months: from May, 2021 to November, 2021. It should also be noted that, as Watts and Stenner (2012) observed, "There is no single or correct way to generate a Q set" (p. 57) and a perfect Q-sample "is probably a thing of fantasy and fiction" (p. 63). This is an important point to consider. In contrast to widely employed in educational research quantitative methodology, known as R methodology, where a standardized instrument is employed, the research instrument in a Q study needs to be custom made in order to suit the context of each particular study. In other words, the Q-sample development − or rather crafting − would be guided not by the standardization demands embedded in R methodology but by weighing nuances of an immediate and specific setting of a Q study, its aims and the research questions it hopes to answer.

CONCLUSION

The main aim of this study was to assess feasibility of adopting Q methodology to explore a complex topic of language-related epistemic beliefs held by Mainland Chinese English language learners. While research on L2 beliefs held by language learners is abundant there is a lack of studies that adopted Q methodology. Considering this gap and recognizing specific features of Q methodology this article highlighted important methodological and theoretical issues that researchers need to be aware of. From a practical application perspective, a particular attention was given to the development of the study concourse and Q-sample. This Q study also considered whether the epistemic beliefs would differ among the language learners at a different level of English language proficiency. A Python code was written to enable a visualization of the findings, which was not done in earlier available Q studies.

The findings from this study provided deeper insights into Chinese L2 learners' subjectively held opinions. They also endorsed the applicability of Q methodology for exploring a complex, multidimensional subject of language-related epistemic beliefs. Moreover, the findings highlighted the prominence within personal epistemologies of the beliefs pertaining to the process of acquiring knowledge of a new language. To conclude, it is hoped that the present study with its detailed account of the concourse and Q-sample development and the insights from piloting the newly-developed instrument would be informative and useful for L2 researchers who wish to conduct their own Q study.

DECLARATION OF COMPETING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Yanyan Wang: investigation, methodology, formal analysis, writing - original draft, writing - review & editing

Larisa Nikitina: conceptualization, methodology, supervision, writing - original draft, writing - review & editing

Jagdish Kaur: supervision, writing - review & editing

Fumitaka Furuoka: formal analysis, software, visualization

REFERENCES

Albright, E.A., Christofferson, K., McCabe, A., & Montgomery, D. (2019). Lessons learned: Some guidelines to factor interpretation. *Operant Subjectivity: The International Journal of Q Methodology, 41*, 134-146. Doi: 10.15133/j.os.2019.001

Banasick, (2019). KADE: A desktop application for Q methodology. *Journal of Open Source Software, 4*(36), 1360. Doi: 10.21105/joss.01360

Barcelos, A. M. F., & Kalaja, P. (2011). Introduction to beliefs about SLA revisited. *System, 39*(3), 281-289. Doi: 10.1016/j.system.2011.07.001

Block, J. (2008). *The Q-sort in character appraisal: Encoding subjective impressions of persons quantitatively.* American Psychological Association.

Brown, S. R. (1980). *Political subjectivity: Applications of Q methodology in political science*. New Haven, CT: Yale University Press.

Brown, S. R. (1993). A primer on Q methodology. *Operant Subjectivity, 16*(3/4), 91-138.

Brown, S.R. (2008). Q methodology. In Lisa M. Given (Ed.), *The SAGE encyclopedia of qualitative research methods* (pp. 699-702). Thousand Oaks, CA: Sage,

Brown, S. R. (2019). Subjectivity in the human sciences. *The Psychological Record, 69*(4), 565-579. Doi: 10.1007/s40732-019-00354-5

Caruso, M., & Fraschini, N. (2021). A Q methodology study into vision of Italian L2 university students: An Australian perspective. *The Modern Language Journal, 105*(2), 552-568. Doi: 10.1111/modl.12713

Chan, K. W., & Elliott, R. G. (2002). Exploratory study of Hong Kong teacher education students' epistemological beliefs: Cultural perspectives and implications on beliefs research. *Contemporary Educational Psychology, 27*(3), 392-414. Doi: 10.1006/ceps.2001.1102

Damio, S. M. (2018). The analytic process of Q methodology. *Asian Journal of University Education (AJUE), 14*(1), 59-75. Retrieved from https://ir.uitm.edu.my/id/eprint/21907/1/AJ_SITI%20MAFTUHAH%20DAMIO%20AJUE%2018.pdf

Damio, S. M., & Hashim, H. (2014). Conceptions of autonomy in language learning from a cultural perspectives using Q methodology. *IIUM Journal of Educational Studies, 2*(1), 33-45. Doi: 10.31436/ijes.v2i1.27

Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.

Fong, E. T. Y. (2021). *English in China: Language, identity and culture*. London: Routledge.

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology.

Contemporary Educational Psychology, 25, 378–405. Doi:10.1006/ceps.1999.1026

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67*(1), 88-140. Doi: 10.3102/00346543067001088

Irie, K. (2014). Q methodology for post-social-turn research in SLA. *Studies in Second Language Learning and Teaching, 4*(1), 13-32. Retrieved from https://files.eric.ed.gov/fulltext/EJ1134810.pdf

Irie, K., & Ryan, S. (2015). Study abroad and the dynamics of change in learner L2 self-concept. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds*.), Motivational dynamics in language learning*. Bristol, UK: Multilingual Matters.

Irie, K., Ryan, S., & Mercer, S. (2018). Using Q methodology to investigate pre-service EFL teachers' mindsets about teaching competences. *Studies in Second Language Learning and Teaching, 8*(3), 575-598. Retrieved from https://pressto.amu.edu.pl/index.php/ssllt/article/ view/9444

Jehng, J. C. J., Johnson, S. D., & Anderson, R. C. (1993). Schooling and students' epistemological beliefs about learning. *Contemporary Educational Psychology*, *18*(1), 23-35. Doi: 10.1006/ceps.1993.1004

Kahsay, M. T. (2019). EFL students' epistemological beliefs and use of cognitive and metacognitive strategies in Bahir Dar University. *International Journal of Foreign Language Teaching and Research, 7*(26), 69-83. Retrieved from http://jfl.iaun.ac.ir/article_625692_feb95289b803bbd11b87451dd5e7fd6a.pdf

Lu, X., Zou, X., & Tao, J. (2020). Spanish language teachers' pedagogical beliefs in Chinese universities. *Círculo de Lingüística Aplicada a la Comunicación, 84*, 41-53. Doi: 10.5209/clac.71994

Lundberg, A. (2019). Teachers' beliefs about multilingualism: findings from Q method research. *Current Issues in Language Planning, 20*(3), 266-283. Doi: 10.1080/14664208.2018.1495373

Lundberg, A. (2020). *Viewpoints about educational language policies: Multilingualism in Sweden and Switzerland*. Doctoral dissertation, Malmö universitet. Retrieved from http://mau.diva-portal.org/smash/get/diva2:1434537/FULLTEXT01.pdf

McKeown, B. F., & Thomas, D. B. (2013). *Q methodology (2nd ed.)*. Los Angeles, CA: Sage.

Mori, Y. (1999). Epistemological beliefs and language learning beliefs: What do language learners believe about their learning?. *Language Learning, 49*(3), 377-415. Doi: 10.1111/0023-8333.00094

Nikitina, L., & Furuoka, F. (2018). Dimensionality in language learners' personal epistemologies. *International Journal of Instruction, 11*(1), 269-282. Retrieved from https://files.eric.ed.gov/fulltext/EJ1165215.pdf

Pajares M. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332. Doi: 10.3102/00346543062003307

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart & Winston.

Pritchard, D. (2006). *What is this thing called knowledge?* Routledge: Abingdon.

Ramlo, S. (2020). Divergent viewpoints about the statistical stage of a mixed method: qualitative versus quantitative orientations. *International Journal of Research & Method in Education*, 43(1), 93-111. Doi: 10.1080/1743727X.2019.1626365

Rock, J. (2013). Exploring vocabulary-related epistemological beliefs with Q-methodology. *L'analisi Linquistica E Letteraria, 21*(2), 137-154. Retrieved from https://www.analisilinguisticaeletteraria.eu/index.php/ojs/article/view/282

Schommer, M. (1998). The influence of age and education on epistemological beliefs. *British Journal of Educational Psychology, 68*(4), 551–562. Doi: 10.1111/j.2044-8279.1998.tb01311.x

Schraw, G., Bendixen, L. D., & Dunkle, M. E. (2002). Development and validation of the Epistemic Belief Inventory (EBI). In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 261–275). Lawrence Erlbaum Associates Publishers.

Stahl, E., & Bromme, R. (2007). The CAEB: An instrument for measuring connotative aspects of

epistemological beliefs. *Learning and Instruction, 17*, 773-785. Doi: 10.1016/j.learninstruc.2007.09.016

Stephenson, W. (1935a). Technique of factor analysis. *Nature, 136*(3434), 297. Doi: 10.1038/136297b0

Stephenson, W. (1935b). Correlating persons instead of tests. *Character and Personality, 4*, 17-24.

Stephenson, W. (1936). The foundations of psychometry: Four factor systems. *Psychometrika, 1*, 195-209.

Stephenson, W. (1953). *The study of behavior: Q-technique and its methodology.* Chicago: University of Chicago Press.

Stephenson, W. (1986a). Protoconcursus: The concourse theory of communication. *Operant Subjectivity, 9*(2), 37-58.

Stephenson, W. (1986b). Protoconcursus: The concourse theory of communication. *Operant Subjectivity, 9*(3), 73-96.

Stephenson, W. (1992). Self in everyday life. *Operant Subjectivity, 15*(2), 29–55. doi:10.15133/j.os.1992.001

Watts, S., & Stenner, P. (2012). *Doing Q methodological research: Theory, method and interpretation*. London, England: Sage. Doi: 10.4135/9781446251911

Wood, P., & Kardash, C. (2002). Critical elements in the design and analysis of studies of epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 231–260). Lawrence Erlbaum Associates Publishers.

Zheng, Y., Lu, X., & Ren, W. (2019). Profiling Chinese university students' motivation to learn multiple languages. *Journal of Multilingual and Multicultural Development, 40*(7), 590-604. Doi: 10.1080/01434632.2019.1571074

APPENDIX

English version of the Q-sample on language learners' epistemic beliefs

1. English language is simpler than language teachers make you think.
2. An English language teacher's job is to give the students answers and not ask them to find the answers for themselves.
3. A good way to learn English language is to re-organize the information according to one's own personal understanding.
4. Connecting new knowledge, such as grammar and vocabulary, with existing knowledge is a good way to learn English.
5. Memorizing vocabulary and grammar is all that is needed to know English.
6. Most English words have one clear meaning.
7. The best thing about English knowledge is that there is always one right answer.
8. English language learners' knowledge of this language can never be certain.
9. The English language teacher's explanations and answers must be exactly the same as in the textbook.
10. Even if a person knows English well, there still remain many things to be discovered about this language.
11. English language never changes. In the future, it will be the same as today.
12. There are no puzzling problems in English grammar.
13. English language is constantly evolving and changing.
14. Having a good knowledge of English means to know lots of grammar rules.
15. When people learn new English grammar, this knowledge is certain and has been agreed upon by linguists and language experts.
16. It is ok to doubt what English language teacher says.
17. I never doubt information about English language use that I receive from native speakers of English.
18. English language teachers are the ones who can help students when they have difficult problems with English.
19. Not every English language teacher has a perfect knowledge of the language they teach.
20. In order to learn a foreign language well, language learners need to be able to distinguish reliable sources of knowledge from unreliable.
21. I do not have to believe everything that native speakers say about English.
22. The most authoritative knowledge of English comes from linguists and language experts.
23. Language learners who disagree with native speakers about grammar or vocabulary usage are over-confident.
24. Even a textbook published by a famous publisher may not give good knowledge of English.
25. Knowledge of English comes from multiple sources rather than one single source.
26. Knowledge of English mainly comes from textbooks.
27. Knowledge of English mainly comes from language teachers.
28. Internet and language apps are reliable sources of English language knowledge.
29. Language knowledge comes from one's own experience of using it rather than from books.
30. To gain knowledge you need to discover how to learn.
31. Some people have a talent for language learning, and others do not.
32. A poor language learner can be trained to learn English well.
33. Good study skills make little difference if you are not naturally good at learning languages.
34. Students who are "average" in learning English will be remain "average" in learning other languages.
35. If someone cannot learn new English vocabulary fast this person will never learn it.
36. Language learning is a slow and gradual process.
37. If a language learner cannot immediately understand a new grammar rule he or she should keep trying to understand it.
38. Language learning is quick. If you cannot learn it fast it is not worth trying.
39. Knowledge of English can't be acquired at once, it needs constant accumulation.
40. If a language learner reads explanations about difficult grammar rules many times, he or she will be able to understand and learn these rules well.
41. People can study English language for years and still not have a good knowledge of it.
42. Achieving high proficiency in English requires a lot of hard work.
43. Effort is overshadowed by talent.
44. Even for a smart student, it takes a lot of perseverance to learn English.
45. Everyone can learn English well if they work hard enough.

https://doi.org/10.17323/jle.2022.12855

Flipped Classroom Model For EFL/ ESL Instruction in Higher Education: A Systematic Literature Review

Dodi Siraj Muamar Zain

Universitas Muhammadiyah Purwokerto, Indonesia

ABSTRACT

Background. Researchers and teaching practitioners are working on formulating instructional approaches that suit the students' interests by incorporating the latest technology. The flipped classroom model emerges as an alternative to extend the classroom interaction, create a different atmosphere, and accommodate a meaningful and collaborative interaction in the class.

Purpose. This paper aims to present a systematic review of past studies focusing on the implementation of the flipped classroom for EFL/ESL instruction in higher education.

Methods. A total of 29 articles were selected using specific criteria comprising year and types of publication, contexts, and the determined quality standard. The collected articles were then analyzed to reveal the research components including the research designs, participants, and instructional tools.

Results. This study presents a discussion of the practical instructional benefits of the flipped classroom model and the possible barriers.

Implications. Furthermore, from this discussion, this study formulated pedagogical and research implications for possible directions in future studies.

KEYWORDS

flipped classroom model, ESL/EFL, systematic literature review

INTRODUCTION

The concept of Flipped Classroom Model (FCM) was first introduced by Lage, Platt, and Treglia (2000) referring to the idea of the inverted version of the traditional classroom (Lage, Platt & Treglia, 2000; Deng, 2019). Since then, this concept has attracted more attention as Bergmann and Sams (Milman, 2012) integrated the latest technology into the concept that was later recognized as FCM. The continuous development of educational technology has triggered various innovations that transform the face of traditional teaching approaches. Nowadays, it has also received an increased research interest in various disciplines along with the emergence of online educational platforms such as Moodle, Khan academy, and Ted (Gaughan, 2014).

FCM reverses what is practically done in a traditional class as learning content is taught outside the classroom while the regular take-home assignments are discussed in the classroom. Zhang (2019) asserted that FCM serves as an innovative teaching approach that incorporates conventional classroom and computer network technology. It can be made possible with technological assistance that enables students to access the learning content from available online learning platforms through internet-connected computers or mobile devices (Jensen, 2019). Additionally, students can be optionally assigned to work in small groups to discuss the content as the instructor monitors their work progress (Bergmann & Sams, 2012). Abeysekera and Dawson (2015) explained that the approach promotes a socially collaborative learning setting representing actual face-to-face learning through digital assistance. Thus, instructors can draw benefits in terms of time management as they can optimize the classroom meeting to discuss pre-delivered learning content (Halili & Zainuddin, 2015, Karami, 2021).

Citation: Zain D. S. M. (2022). Flipped Classroom Model For EFL/ESL Instruction in Higher Education: A Systematic Literature Review. *Journal of Language and Education, 8*(3), 133-149. https://jle.hse.ru/article/view/12855

Correspondence:
Dodi Siraj Muamar Zain,
dodizain@uark.edu

Received: August 10, 2021
Accepted: July 18, 2022
Published: September 30, 2022

Along with the increased popularity of the approach, English as a Foreign Language/ English as a Second Language (EFL/ESL) teaching practitioners and researchers are now crafting a way to incorporate this approach for effective EFL/ESL instructional practices. Previous studies were carried out with a primary focus on the development of students' linguistics skills (e.g. Wu et al., 2017; Lin et al., 2018) while others highlighted the effects of this approach on students' learning motivation (e.g. Yu & Yu, 2017; Cabi, 2018), self-directed learning (e.g. Zainuddin et al., 2019), and students' learning perception (e.g. Hung, 2017; Chen Hsieh et al., 2017). These results attributed to the practicability and potential of this approach confirming a broader extent of learning engagement for enhancing students' subject-matter EFL/ESL learning.

Currently, various resources from the internet such as blogs, vlogs, notes, and briefs have been mainly used as the main references for applying FCM, especially for EFL/ESL contexts. Only a limited number of scientific papers are serving as a reference for an organized and systematic application of this approach. Accordingly, this paper aims to review the previous studies of FCM with an in-depth investigation of the key elements that determine the effectiveness of this approach. The following questions are addressed to cover an extensive range of literature dealing with the research topic and to serve as the outline of this review: (1) What is the theoretical framework used to design FCM? (2) What are the research methodologies and technological tools used to conduct studies of FCM? (3) What are the practical instructional benefits of FCM? (4) What are the barriers to the implementation of FCM? (5) What are the pedagogical research implications of FCM based on previous studies?

METHODS

This study adopts a systematic literature review to examine articles addressing the application of FCM in EFL/ESL contexts. According to Liberati et al. (2009), Systematic reviews are conducted to provide answers to particular scientific questions on a specific theme by collecting and examining the related empirical studies which are aligned with the criteria of the study. To achieve reliable results and analyses, these types of study explicitly mention the adopted methodology and follow particular scientific values (Higgins & Green, 2008; Liberati et al., 2009).

Databases

For this study, the most referred databases in the field of social science and educational technology were employed comprising JSTOR, ProQuest, ERIC, and Google Scholar. This search led to some most preferred instructional technology-related and EFL/ESL teaching journals. The articles were initially collected using specific keywords comprising "*flipped classroom*", "*ESL", "EFL",* and *"higher education"* which revealed various articles with diverse research schemes.

Inclusion and Exclusion Criteria

These initially retrieved articles were then further selected based on specific criteria to address the quality of this review. For instance, the articles needed to be published in the English language and peer-reviewed. The articles should also be based on empirical studies focusing on EFL/ESL instruction in higher education. To address the publication quality, the included articles will only be collected from SCOPUS-indexed journals. Also, to address the novelty of the study, this review only selected published articles in the last six years (2017-2022).

Table 1

Inclusion and exclusion criteria

Criteria	Inclusion	Exclusion
Year of publication	2017 – 2022	Before 2017
Article types	Peer-reviewed	Non-peer reviewed
Types of study	Empirical	Non-empirical
Context of study	EFL/ESL context	Non-EFL/ESL context
Language	English	Non-English
Quality	SCOPUS-indexed	Non-SCOPUS-indexed

The article selection process

The articles used in this review were only limited to empirical study articles. A total of 29 empirical study articles that suit the main topic of this review were collected and thoroughly examined. An article published before 2017 (e.g., Şengel, 2016) was also used as a reference to compare the findings and provide a justifiable interpretation of the analysis and pedagogical implication. In addition, there is inclusion (Vuong et al., 2018) from a non-Scopus journal in this review considering the significance of its findings to support the discussion. The following figure presents the summary of the article selection process.

A total of 29 articles were selected based on the inclusion and exclusion criteria. The majority of them were published in educational technology-related journals while some others were collected from education-related journals. The distribution of reviewed articles based on the referred journal is presented in table 2.

Figure 1

Overview of the paper selection process

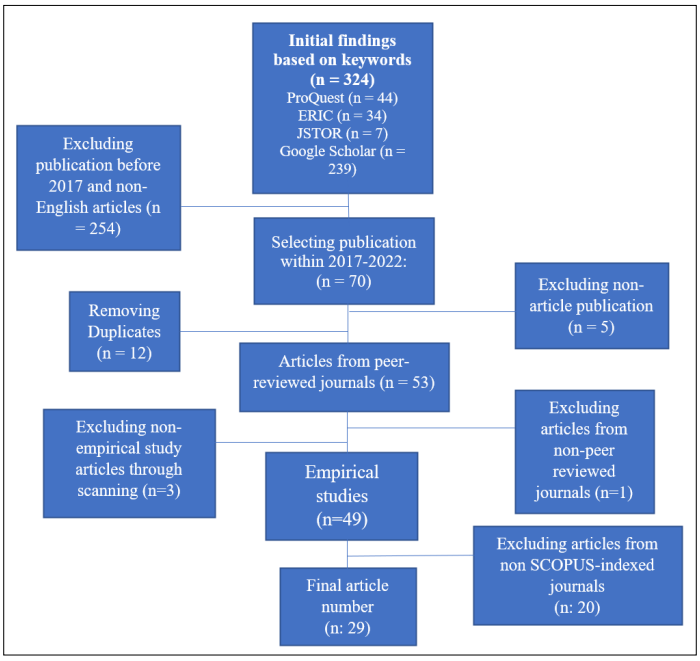


Table 2

Distribution of reviewed studies based on journal references

Journal Reference	Number of studies	Study
International Journal of Emerging Technologies in Learning	7	Abdullah et al. (2019A); Aprianto et al. (2020); Li and Qu (2019); Liu and Zhang (2018); Santikarn and Wichadee (2018); Zhang (2019); Abdullah et al. (2021)
Sustainability	1	Li and Li (2022)
Education and Information Technologies	1	Abdullah et al. (2019B)
Computers and Education	1	Bakla (2018)
Journal of Asia TEFL	4	Bonyadi (2018); Teng (2017); Chen and Liu (2019); Yang et al. (2019)
International Review of Research in Open and Distance Learning	1	Cabi (2018)
Computers in Human Behavior	1	Chen Hsieh et al. (2017)
TESOL Quarterly	1	Lee and Wallace (2018)
The EUROCALL Review	1	Leis and Brown (2018)
Educational Technology and Society	2	Lin and Hwang (2018); Lin et al. (2018)
Smart Learning Environments	2	Öztürk and Ünal (2021); Öztürk and Çakıroğlu (2021)
Journal of Computer Assisted Learning	1	Shyr and Chen (2018)
CALL-EJ	1	Vaezi et al. (2019)
Journal of Educational Technology & Society	1	Wu et al. (2017)
Indonesian Journal of Applied Linguistics	1	Zainuddin et al. (2019)
PLoS ONE	1	Zhonggen (2019)
Technology, Knowledge, and Learning	1	Akayoğlu (2021)
Sage Open	1	Yu and Gao (2022)

RESULTS

The following section presents the analysis of articles to provide answers to the proposed research questions. The analysis revealed the matrix of the latest FCM studies, practical instructional benefits, barriers to implementation, and pedagogical implications of the approach.

Matrix of the Current Study

Using a holistic approach, the matrix of the previous studies presents the overview of previous studies for FCM application in the following table.

Table 3

Matrix of the previous studies

Study	Topic	Theoretical/ pedagogical Framework	Participants	Designs, instruments	Instructional tools
Chen Hsieh et al. (2017)	EFL students' technological acceptance	The Technology Acceptance Model (TAM)	42 sophomores in central Taiwan whose ages ranged from 20 to 21 years old	Action research coupled with a mixed-method approach using tests of oral proficiency, questionnaires, interviews	LINE
Zhonggen (2019)	Video lecturing's effects on student satisfaction and English proficiency	Constructivism	87 Chinese university students, upper intermediate level (Male n = 43, Female n = 44) They ranged from 19 to 22 in age (M = 20.33, SD = 1.02).	Experimental study coupled with a mixed-method approach using Four scales, including a CET 4 to determine English proficiency and three scales to identify interaction feasibility, self-efficacy, and self-regulation levels	Clicker
Cabi (2018)	The effects of FCM on students' academic achievement	Constructivism	59 pre-service teachers studying English Language Teaching and 59 students in Turkish Language Teaching Programs	Experimental study coupled with a mixed-method approach employing achievement tests and focus group interview	Video, Kahoot
Bonyadi (2018)	The impacts of the flipped instruction on students' oral interpretation performance	Discovery learning, socio constructivism	39 Iranian EFL students at the advanced level majoring in English Translation Studies at Islamic Azad University, Urmia branch.	Quasi-experimental design using tests of students' oral interpretation performance	Learning content from various websites
Yang et al. (2019)	Comparing high- and low-achievers English vocabulary learning in a flipped classroom setting	Constructivism, F-L-I-P™ principle	87 second-year students from a university in Northern Taiwan	Experimental design using An English vocabulary assessment, and a questionnaire	Video, Facebook, Kahoot
Leis and Brown (2018)	The effects of instructors' experience in the flipped classroom model	N/A	38 second-year students of a university in Japan	Quasi-experimental design employing a writing test	Videos

Study	Topic	Theoretical/ pedagogical Framework	Participants	Designs, instruments	Instructional tools
Bakla (2018)	Examining student-generated materials in a flipped pronunciation class	Inquiry-based learning	40 intermediate Turkish students of English major students (35 females and 5 males. 34 of the participants (85%) were aged between 18 and 20.	Case study combined with sequential explanatory mixed-methods approach using a post-instruction survey and follow-up interviews	Videos, Moodle, online text-to-speech software, HotPotatoes (Half-baked Software, n.d.) and Softchalk (Softchalk LLC., 2015), online flashcard programs
Vaezi et al. (2019)	Teaching listening comprehension using FCM	Constructivism	119 advanced level students majoring in English translation, literature, and English teaching at two branches of Islamic Azad University	Quasi-Experimental design with a quantitative approach employing a pretest, a posttest, and a delayed posttest	Audio materials, Video and audio resources from TED (www.ted.com), YouTube, and National Public Radio (www.npr.org).
Abdullah et al. (2019A)	The effectiveness of FCM on English speaking performance	Socio-constructivism	27 undergraduate students at intermediate to upper-intermediate level majoring in English at Buraimi University College (BUC), Oman	Quasi-experimental design coupled with A mixed-method employing approach oral proficiency tests, observation notes, focus group interview	Google Classroom
Zainuddin et al. (2019)	Self-directed learning in the EFL flipped-class pedagogy	Constructivism	10 undergraduate students at the intermediate level of a university in Indonesia	Case study adopting a qualitative approach employing online observations and semi-structured interview	TES BlendSpace
Shyr and Chen (2018)	Students' self-directed learning and language performance	Socio-constructivism	81 sophomores of non-English majors at a public university in central Taiwan. Their ages ranged from 19 to 23.	Quasi-experimental study (quantitative) employing a learning performance test, weekly quizzes, a flipped classroom readiness survey, and a self-regulated survey	A technology□enhanced flipped language learning system (Flip2Learn)
Lin and Hwang (2018)	Influencing factors of EFL students' oral performance	Online community-based flipped learning approach	49 EFL freshmen in a Taiwanese university who had the same level of English proficiency and whose age mean was 18	Quasi-experimental design (quantitative) using oral performance video clips, students' participation, students' interactive behavior, and a survey questionnaire	Facebook
Lee and Wallace (2018)	Students' learning outcomes and perceptions	N/A	79 EFL students at intermediate-level in a college English class at a South Korean university. (61 males and 18 females)	Action research design coupled with a mixed-method approach using students' achievements in three major tasks, three surveys, and observation notes	Video lectures

Study	Topic	Theoretical/ pedagogical Framework	Participants	Designs, instruments	Instructional tools
Lin et al. (2018)	A flipped contextual game-based learning approach to improve writing skills	A flipped contextual game-based learning approach, constructivism	68 Taiwanese undergraduates who had a similar level of English proficiency; The students' ages ranged from 18 to 22	Quasi-experimental study coupled with a mixed-method approach, design collecting students' essays, participation, reflections, and online learning behavior as instruments of the study	A contextual educational game
Liu and Zhang (2018)	Computer-Distance Education (CDE) to teach English writing	Constructivism and humanistic learning theories	400 college students among four universities with the course of practical English writing in China	Case study using questionnaires followed by a quasi-experimental study using tests	Jukuu (web-based teaching platform)
Li and Qu (2019)	Developing a platform for teaching English at the college level	Humanistic learning theory, Constructivist theory, Autonomic, and cooperative learning theories	45 students in the non-English majors at the sophomore level in a university in Hunan Province, China	Quasi-Experimental design coupled with a mixed-method approach employing an interview, a questionnaire survey, and TOEIC	The flipped classroom-based B/S teaching platform
Zhang (2019)	Developing an FC-based education system for teaching English at the college level	N/A	152 sophomore students from non-English majors in a university in Zhejiang, China	Case study coupled with a quantitative approach employing questionnaires	A college English education system developed using SQL server and FC.
Santikarn and Wichadee (2018)	FCM to teach EFL at the college level	N/A	40 students enrolled in an advanced English course in a private university in Thailand	Case study coupled with a quantitative approach using five assignments and two sets of questionnaires	Edmodo
Aprianto et al. (2020)	Investigating the effects of multimedia-assisted FCM on students' self-directed learning	N/A	15 undergraduate students of the English education department in an Indonesian University	Case study coupled with a qualitative approach employing observation, questionnaires, and interviews	Various online reading materials and videos learning management systems (LMS)
Abdullah et al. (2019B)	The effects of FCM on students' motivation level in English speaking performance	Constructivism	27 students majoring in English at Buraimi University College (BUC), Oman	Quasi-experimental study coupled with a mixed-method approach employing a Motivation in English Speaking Performance Questionnaire (MESPQ) and reflective journal forms	Instructional videos
Chen and Liu (2019)	Examining high and low achievers' preferences of workload, work type, and perceived usefulness in FCM	Constructivism	39 university foreign language (FL) learners with diverse achievement profiles from south Taiwan, including eight males and 31 females (mean age: 21.5 years)	Case study coupled with a mixed-method approach using questionnaire survey and oral interview	An interactive class platform, Zuvio that includes recorded video lectures
Wu et al. (2017)	Creating an online learning community in a flipped classroom to enhance EFL learners? oral proficiency.	The Community of Inquiry (CoI) framework	50 English-majored sophomores enrolled in two oral training classes at a four-year comprehensive university in central Taiwan	Case study coupled with a mixed-method approach employing oral reading and comprehension tests, a CoI questionnaire, and semi-structured focus-group interviews	Mobile application LINE

Study	Topic	Theoretical/ pedagogical Framework	Participants	Designs, instruments	Instructional tools
Öztürk and Ünal (2021)	FCM as a self-regulated learning strategy to develop language skills	N/A	49 students (18–20 years old) enrolled in an EFL course at a state university accounting department in Turkiye	Quasi-experimental pre-test-posttest control group design (quantitative) using English proficiency tests	FCM_WEB platform to access instructional videos, reading notes, and online test
Teng (2017)	The effects of FCM on the improvement of students' academic performance and learning satisfaction	Constructivism, Problem-based theories, and Self-determination theory	90 students learning English as a foreign language (EFL) 50 females and 40 males—aged 18-20 years old at a university in China	Experimental design coupled with a mixed-method approach using academic performance tests, questionnaires, and semi-structured interviews	Instructional online videos. (WebQuests plus online videos)
Öztürk and Çakıroğlu (2021)	Employing self-regulated learning strategies within a scheme of FCM to improve students' language performance	Self-regulated learning	49 EFL undergraduates at a state university in Turkiye. The students' ages ranged from 18 to 20 years old.	A quasi-experimental pre-test-posttest control group design employing an English proficiency test and an achievement test	FCM_SRL system presenting videos, online tests, online discussion, and diary module
Abdullah et al. (2021)	Examining students' anxiety in the production of oral language under a scheme of FCM	Collaborative and active learning approaches	27 undergraduates majoring in English as a foreign language at Buraimi University College (BUC), Oman.	Case study, a mixed-method approach employing the Anxiety in English Speaking Performance Questionnaire (AESPQ), the focus group interviews, and students' reflective journals	Video lectures
Akayoğlu (2021)	The perceptions of EFL pre-service teachers on FCM	Self-determination theory	58 pre-service teachers of English in Turkiye consisting of 43 females and 15 males. The ages range from 21 to 24 years old	Case study and a mixed-method approach using the questionnaires and interviews	Google Classroom
Yu and Gao (2022)	The effects of video length on English proficiency, student engagement, and satisfaction in a flipped English classroom	Constructivism	97 Chinese university students (Female n = 75; Male n = 22) The ages range from 16 to 19 years old (M = 17.49; SD = .98).	Experimental design coupled with a quantitative approach using The TOEFL test and questionnaires	Instructional videos
Li and Li (2022)	The effects of the flipped of FCM on four aspects of learner engagement in teaching listening and speaking	Collaborative learning	69 university students in mainland China	Experimental design coupled with a mixed-method approach using tests and semi-structured interviews	The Cloud Classroom, Instructional videos

Table 3 above presents a brief description of the studies comprising the topic, the conceptual/theoretical frameworks, participants, research design and instruments, and the instructional tools. In the following section, the results and discussions of different components of studies are presented exposing the pedagogical benefits, practical challenges, and the formulated pedagogical and research implication.

Theoretical Frameworks in the Previous Studies

The majority of the studies (n = 12) were grounded on the scheme of constructivism (Yu & Gao, 2022; Zhonggen, 2019; Cabi, 2018; Bonyadi, 2018; Yang et al., 2019; Vaezi et al. 2019; Zainuddin et al., 2019; Abdullah et al., 2019B, Chen & Liu, 2019; Lin et al., 2018; Liu & Zhang, 2018; Li & Qu, 2019). Constructivists believe that learners build their own knowledge by integrating new information with their pre-existing knowledge instead of obtaining information passively. Three other studies by Abdullah et al. (2019A), Bonyadi (2018), and Shyr and Chen (2018) indicated the application of Socio constructivism. This theory postulates the essence of interaction as a pivotal element of students' learning. Based on this theory, learning occurs as ideas are exchanged and constructed via interaction between the instructor and students or among students (Powell & Kalina, 2009). Bonyadi (2018) also included discovery learning which was introduced by Jerome Bruner (1961). This popular theory highlights the idea that learners should be directed to acquire new information and truths by building on past experiences, pre-existing notions, imagination, and creativity.

Other studies were based on various schemes of theories. For instance, a study by Yang et al. (2019) was conducted based on the F-L-I-P™ principle. This principle emphasizes the need for flexibility in the learning environment, self-directness in learning, qualified learning materials, and the professionalism of the educators. Aside from constructivism, Both Li and Qu (2019) and also Liu and Zhang (2018) grounded their study on humanistic learning theory. This theory postulates that the learning process occurs as feeling intervenes with knowledge acquisition. It emphasizes the essence of both cognitive and affective learning as the determinants of the success of learning (Chen et al., 2017). Wu et al. (2017) implemented The Community of Inquiry (CoI) framework which serves as a model to examine and evaluate a learning community (Garrison et al., 2001; Garrison et al., 2010). CoI incorporates three interdependent components to manifest effective online learning: teaching presence, social presence, and cognitive presence.

Furthermore, Li and Qu (2019) developed their study based on the autonomic theory and the cooperative learning theory. The autonomic learning theory explains the importance of the learners' subjective initiative to internally shape learning behavior. The cooperative learning theory focuses on the use of students' grouping to complete classroom assignments or discuss particular issues presented in class. Inquiry-based learning was adopted in Bakla (2018). It refers to a learning model in which students generate questions, formulate answers to respond to posed questions, communicate the answers and new information with their peers, and do a reflection on the learning process. Teng (2017) and Akayoğlu (2021) grounded on self-determination theory to examine the effects of FCM on the improvement of students' linguistic performance and learning satisfaction. This theory postulates the division of students' learning motivation into two categories: intrinsic and extrinsic motivation (Deci & Ryan, 2002). Intrinsic motivation is related to the emotional state as a result of students' engagement in particular activities. Extrinsic motivation is related to the external factors which empower students for their involvement in particular activities.

Öztürk and Çakıroğlu (2021) adopted Self-Regulated Learning (SRL) which is based on the work of Zimmerman (1990). This framework emphasizes the significance of the students' metacognition, motivation, and behavior as supporting elements for successful learning. Students have the authority to manage their own learning through a cyclical process of SRL comprising three phases of learning; forethought phase, performance phase, and self-reflection phase (Öztürk & Çakıroğlu, 2021; Artino & Stephens, 2009; Zimmerman, 1998).

Lastly, Li and Li (2022) adopted a collaborative learning model. This model emphasizes the essence of students' collaboration through learning groups. Each group member is expected to contribute to the group discussion by sharing knowledge and experience and encouraging other group members to achieve the goals of learning (Johnson & Johnson, 1986).

Participants of the Study

All studies in this review involved EFL/ESL students of higher education from different parts of Asia nations. For instance, a total of 16 studies were conducted in East Asia such as China (Yu & Gao, 2022; Li & Li, 2022; Zhonggen, 2019; Zhang, 2019; Teng, 2017; Liu & Zhang, 2018; Li and Qu, 2019), Japan (Leis & Brown, 2018), Taiwan (Yang et al., 2019; Shyr & Chen, 2018; Chen & Liu, 2019, Lin & Hwang, 2018; Lin et al., 2018; Wu et al., 2017; Chen Hsieh et al., 2017) and South Korea (Lee & Wallace, 2018). In Southeast Asia, two studies were reported from Indonesia (Zainuddin et al., 2019; Aprianto et al., 2020) and one study was conducted in Thailand (Santikarn & Wichadee, 2018). In Western Asia, two studies were conducted in Iran (Vaezi et al., 2019; Bonyadi, 2018) and the other three studies were from Oman (Abdullah et al., 2019A; Abdullah et al., 2019B; Abdullah et al., 2021). This study also included five studies from Turkiye (Cabi, 2018; Bakla, 2018; Öztürk & Ünal, 2021; Öztürk & Çakıroğlu, 2021; Akayoğlu, 2021). Overall, a total of 2032 participants ranging from 18 to 23 years old were recruited in 29 studies.

Methods and Instruments

The majority of the studies (n = 12) in this review adopted the quasi-experimental design coupled with either a quantitative approach (n = 8) or a mixed-method approach (n= 4). For example, Bonyadi (2018) employed a quasi-exper-

imental design that was quantitatively structured through students' oral interpretation performance tests as data collection. Similarly, Leis and Brown (2018) employed this design to examine the effects of instructors' experience in the flipped classroom model and collected the data through writing tests. A combination of a quasi-experimental design and mixed-method approach can be found in the study by Abdullah et al. (2019A) which employed oral proficiency tests, observation notes, and focus group interviews as the instruments of data collection.

Aside from quasi-experimental design, six studies included randomization in subject selection to incorporate an experimental design (Li & Li, 2022; Yu & Gao, 2022; Zhonggen, 2019; Cabi, 2018; Yang et al., 2019; Teng, 2017). Zhonggen (2019) adopted an experimental design combined with a mixed-method approach using four scales, including a CET 4 to determine English proficiency and three scales to identify interaction feasibility, self-efficacy, and self-regulation levels. Also, Cabi (2018) combined an experimental design and a mixed-method approach to gather data through achievement tests and focus group interviews.

A total of 10 studies were identified using a case study which was coupled with either a qualitative (Zainuddin, Habiburrahim, Muluk, & Keumala, 2019; Aprianto, Purwati, & Anam, 2020), a quantitative (Zhang, 2019; Santikarn & Wichadee, 2018), or a mixed-method approach (Bakla, 2018; Chen & Liu, 2019; Wu et al., 2017; Abdullah et al., 2021; Akayoğlu, 2021). Zainuddin et al. (2019) employed a case study design and qualitatively collected data through approach online observations and semi-structured interviews. Meanwhile, Aprianto et al. (2020) combined a case study design and a mixed-method approach by employing observation, questionnaires, and interviews. Both Zhang (2019) and Santikarn and Wichadee (2018) adopted a case study and collected data through questionnaire surveys with additional instruments of five assignments for Zhang (2019). Additionally, a study by Liu and Zhang (2018) combined a case study design using questionnaires and a quasi-experimental design using tests.

Lastly, two studies adopted an action research design coupled with a mixed-method approach. Lee and Wallace (2018) adopted this combination and harnessed three types of instruments comprising students' achievements in three major tasks, three surveys, and observation notes. Meanwhile, Chen Hsieh et al. (2017) also used three instruments including tests of oral proficiency, questionnaires, and interviews.

Instructional Tools

Several different tools were employed in the reviewed studies. They include social media, instructional platforms, multimedia content, and other software/tools.

Social Media

Facebook has been a popular social media to support and promote student self-learning. In this review, two studies adopted this social media as a tool to bridge interaction among students and between the instructor and students. Yang et al. (2019) used Facebook as one of the tools in their study that compared high- and low-competent students regarding their skills in English vocabulary. Meanwhile, Lin and Hwang (2018) utilized this media as a primary tool to examine the influencing factors of EFL students' oral performance. On the Facebook wall, the instructor provided several instructional video links retrieved from TED, Voice Tube, and YouTube. Mobile application LINE which expands its function from a mobile messaging application to a mobile social media application was used in both Chen Hsieh et al. (2017) and Wu et al. (2017).

Instructional Platforms

Several studies mentioned the use of instructional platforms for the application of FCM. This review noted various platforms used in different schemes of studies that are available in the market comprising google classroom (Abdullah et al., 2019A; Akayoğlu, 2021), Web-based teaching platform Jukuu (Liu & Zhang, 2018), Moodle (Bakla, 2018), Edmodo (Santikarn & Wichadee, 2018) and an interactive class platform, Zuvio (Chen & Liu, 2019). In a study, Zainuddin et al. (2019) utilized TES BlendSpace to develop the lesson by posting instructional video lessons retrieved from YouTube and arranging class assignments.

Additionally, several studies claimed to have developed their instructional platform which is specifically tested for the study. For instance, Li and Qu (2019) developed the flipped classroom-based B/S teaching system to present students with pre-class learning content. This system enables students to access various online resources including NetEase Open Class, U.S National English Corpus, China National Knowledge Network, and other learning content designated for students at the higher education level. Similarly, Zhang (2019) also harnessed a college English education system that was constructed based on SQL server and FC. Both Aprianto et al. (2020) and Li and Li (2022) used a particular learning management system (LMS) to share learning content and enable learning collaborations among students and between the instructor and students. In another study, Shyr & Chen (2018) employed A technology-enhanced flipped language learning system (Flip2Learn) which was constructed and developed using Microsoft SQL Server 2005 (the database structure of Flip2Learn). This system collected data about their learning preferences and perspectives as students initially accessed it. Additionally, both Öztürk and Ünal (2021) and Öztürk and Çakıroğlu (2021) developed an FCM_WEB platform to access instructional videos, reading notes, and online tests.

Multimedia Content

To support student self-learning, the instructors used multimedia formats of learning content. It was found that the majority of the studies used videos that were linked to the learning management system. For instance, Vaezi et al. (2019) presented videos and audio content that were linked to TED, YouTube, and National Public Radio. Aprianto et al. (2020) provided the links to online reading references and videos from Youtube on the developed learning management system. Similarly, in Bonyadi (2018), learning content from various websites was made available for students to access during their self-study.

Other videos were specifically created by recording the lectures within a specific time duration. This type of media was found in eight studies (Abdullah et al., 2019B; Cabi, 2018; Yang et al., 2019, Leis & Brown, 2018; Lee & Wallace, 2018; Teng 2017; Abdullah et al., 2021; Bakla, 2018).

Other Software / Tools

Aside from social media, instructional platforms, and multimedia content, several studies included other software or tools to support the FCM application. Cabi (2018) and Yang et al. (2019) used Kahoot! which is a learning platform that is specifically designed for generating, sharing, and playing learning games or short quizzes. Bakla (2018) utilized various software such as online text-to-speech software, HotPotatoes (Half-baked Software, n.d.), Softchalk (Softchalk LLC., 2015), and online flashcard programs.

Incorporating the online gaming elements, Lin et al. (2018) utilized RPG Maker VX from Enterbrain Incorporated to develop a contextual educational game. This game embedded a story-based learning module along with a game rules module to appropriately follow business trade cycles or public regulations. From this game, students can gain a good comprehension of business trade and improve their writing skills in the field of business writing. Additionally, Zhonggen (2019) used a Clicker which refers to a hand-held device to remotely control projected visual output from computers. This device allows students to get engaged in classroom discussion with a better learning efficiency.

Practical Instructional Benefits of the Flipped Classroom Model

The studies of FCM have provided evidence of instructional benefits that positively affect students' competence in EFL instruction. Based on the analysis of the studies, this study reported that FCM contributes to students' learning progress by supporting linguistic skills development, improving learning motivation, encouraging self-directed learning, promoting self-efficacy, and facilitating self-directed learning.

Language Skills Development

FCM is deemed as an innovative approach to teach a target language by incorporating various instructional technology. Various studies have reported the affordances of this approach to develop students' language skills. This review identified four studies (Abdullah et al., 2019; Bonyadi, 2018; Wu et al., 2017; Lin & Hwang, 2018) that demonstrate the positive effects of FCM on students' oral performance. Three studies (Abdullah et al., 2019; Bonyadi, 2018; Lin & Hwang, 2018) employed a quasi-experimental design and shared similar findings despite a difference in the technological tools for the experiment. Based on the findings, all these three studies concluded that the experiment managed to enhance students' oral performance as indicated by the statistical analysis. An interesting finding is found in Lin and Hwang (2018) stating that students with high achievement exhibited higher involvement during the online meeting in comparison with students with low achievement. Meanwhile, Wu et al. (2017), as well as Li and Li (2022), revealed the significance of the online learning community to accommodate engaging and active learning collaboration which subsequently led to the development of students' oral language performance. Information about the potential of FCM to improve students' writing skills was extracted from Lin et al. (2018) and Liu and Zhang (2018). Despite the use of different instructional tools in their FCM applications, both studies implied that FCM was effective to improve the students' English writing skills. The studies indicated a lower quantity of writing errors due to the effect of FCM.

This review included a study emphasizing the effects of FCM on the improvement of reading skills by Lee and Wallace (2018). This study compared a Communicative Language Teaching (CLT) approach with FCM. Both approaches managed to improve students' reading performance. However, the majority of students in FCM group positively perceived the model and exhibited more active participation during the classroom discussion in comparison to the CLT group.

Focusing on listening skills, Vaezi et al. (2019) investigated the effects of FCM on 119 English language learners at the advanced level. This experimental study compared two experimental groups; Authentic Audio Material Group (AAMG), Pedagogical Audio Material Group (PAMG), and one Control Group (CG). The results revealed that the treatment through FCM for Authentic Audio Material Group (AAMG) and Pedagogical Audio Material Group (PAMG) managed to effectively improve students' listening comprehension in both the short and long term. Both of these groups managed to outperform the control group. However, the authors indicated more positive effects of authentic audio content compared to pedagogical audio content.

Other studies indicated that the FCM application affects students' academic progress as seen in their classroom performance. Li and Qu (2019) compared the experiment group

receiving FCM treatment that adopted B/S architecture platform with the control group that was taught using the conventional model. The study implied the effectiveness of FCM to improve students' reading and listening skills as indicated in the comparison of English test scores between both groups. Substantially, the increase in the English scores for the experimental group was higher than that of the control group. In another study, Yang et al. (2019) argued that this instructional approach managed to develop vocabulary mastery with an indication that low achievers gained more benefits from this approach compared to high achievers.

Learning Motivation

Improving students' learning motivation serves as an appeal of FCM to many teaching practitioners in the current years. The English proficiency of college students is primarily reliant on the level of their learning motivation. This review identified five studies that explain the impact of FCM on students' learning motivation. Teng (2017) demonstrated how FCM improves students' positive perceptions of learning during their involvement in a cross-cultural communication course in an experimental study. Another experimental study by Cabi (2018) compared flipped-based blended learning and conventional blended learning. Both treatments did not show any significant difference, but students were more motivated to learn in flipped-based blended learning due to the absence of homework. Meanwhile, Zhang (2019) developed a particular system of FCM based on SQL Server with the browser/server (B/S) structure. The system managed to encourage students to learn the content independently which subsequently increase their time for self-study. Students were more motivated to participate during class discussions and thus developed their skills and knowledge in learning EFL. Two other studies by Abdullah et al. (2019B) and Abdullah et al. (2021) revealed that students became more productive in language production and less reluctant to express their ideas during classroom discussions. Also, FCM allows them to use their own preferred learning strategies to complete their class assignments.

Students' motivation during FCM is affected by several factors. A study by Li and Li (2022) addressed this issue and found that the learning environment, instructor presence, learning content, and learner presence contributed to the improvement of students' motivation to actively participate in class discussions. More specifically, Yu and Gao (2022) argued that FCM could elevate students' motivation by considering particular elements including the length of learning videos and the flexibility of the communication styles between students and the instructor.

Self-directed Learning

FCM is reported to accommodate self-directed learning among EFL/ESL students. This review found nine studies explaining the strengths of FCM to accommodate self-directed learning. For example, Santikarn and Wichadee (2018) and Aprianto et al. (2020) claimed that the application of FCM allowed students to be more independent in their language learning. FCM could assist students in independently managing their learning and become more creative to gain comprehension of the learning content during the learning process while lowering the learning tension. This claim is supported by Zhonggen (2019) and Bonyadi (2018) who reported that FCM can be a suitable approach for students who prefer to be more independent in learning and not reliant on their instructors' guidance to study the learning content and solve learning issues. Students can make the most of their learning by harnessing their respective learning styles (Leis & Brown, 2018; Shyr & Chen, 2018; Lin & Hwang, 2018). In more recent studies, Öztürk and Ünal (2021) and Öztürk and Çakıroğlu (2021) provided evidence that both FCM and self-directed learning strategies could be an effective combination to improve students' language skills.

Self-Efficacy

Several studies provided evidence of the positive impacts of FCM in developing students' self-efficacy in English language learning. For instance, Liu (2019) asserted that this model brings refreshment to the conventional instruction of EFL. Through this model, students were positively triggered to actively engage in the learning interaction and could easily adapt to the transition from conventional face-to-face instruction to FCM (Aprianto et al., 2020; Li & Qu, 2019; Zhonggen, 2019). Students were also more confident with their language performance during the learning process (Abdullah et al., 2019A; Chen-Hsieh et al., 2017). FCM also provides more engaging and attractive learning activities which subsequently led to students' learning satisfaction and self-confidence in using technologies (Lee & Wallace, 2018; Lin & Hwang, 2018; Santikarn & Wichadee, 2018; Akayoğlu, 2021). Furthermore, Abdullah et al. (2019B) further indicated the development of positive attitudes toward FCM among students. It was reported that students were satisfied with the class due to increased opportunities for them to practice their verbal communication skills.

Collaborative Learning

FCM is incorporated to meet the needs of EFL students to adapt to technological development and prepare themselves for engaging and collaborative instructional methods (Dorman, 2016). This approach facilitates a learning setting in which students can collaboratively work to solve emerging issues or gain the required notion (Bonyadi, 2018; Zhonggen, 2019). Aprianto et al. (2020) and Zainuddin et al. (2019) stated that FCM promoted better learning involvement, interaction, and collaboration among the students. Students could help other students to check their peers' errors of language production while, at the same time, becoming more aware of their abilities and mistakes during online peer assessment. Moreover, Chen and Liu (2019) explained

the different preferences of learning settings as highly competent students like to work individually while low competent students like to work collaboratively in the FCM setting.

Barriers to the Implementation

Aside from the aforementioned pedagogical benefits of the FCM as an effort to develop students' EFL language performance and provide an engaging learning experience, there are some indications of the barriers to the effective implementation of this approach. These barriers include the problems of self-directed learning, technical issues, and increased workload.

Problems of Self-directed Learning

One strong point of FCM is that this model allows students to be independent to manage their own learning. However, students might find it difficult to concentrate during the learning process due to the distraction of their surroundings (Vuong et al., 2018). Cabi (2018) further asserted that some students still find it difficult to manage their time for studying.

Technical Issues

To effectively utilize FCM, both students and instructors need to have proper access to the required technology. Vuong et al. (2018) indicated students' frustration as a result of a slow internet connection indicating the need for qualified internet-connected devices and a good internet connection. A similar issue was found in Li and Li (2022) which also addressed the negative impacts of using long videos to present the learning content. In another study, Zhonggen (2019) reported that video lecturing impeded students' reading comprehension as this medium often contains some distracting noise while being played. Eye fatigue also became a reason why students preferred to read learning content from hardcopy files instead of mobile content. Additionally, videos or PowerPoint slides might restrict the instructor's opportunities to immediately clarify the unclear content (Zhonggen, 2019). The quality of tools used in a flipped classroom was also a concern in a study by Teng (2017) and Cabi (2018). The use of low-quality videos to substitute class lectures posed learning difficulties for students to comprehend the learning content.

Increased Workload

FCM allows an extensive classroom discussion that requires instructors to present learning contents in various format and suitable teaching strategies. It can be a challenging task for instructors as they need to allocate their time to create the materials or collect them from third parties that are suitable for students' needs (Zhonggen, 2019). Suranakkharin (2017) further suggested that instructors need to train themselves to be proficient in organizing classroom activities as well as harnessing online learning facilities. It is not an easy task, especially for those who are not accustomed to online technology. Students, on the other hand, made extra efforts in the flipped classroom. Bakla (2018) mentioned the need for various digital literacies which often require certain preparation to implement the model. Furthermore, FCM poses students with additional workloads that might be unappealing and even stressful (Vuong et al., 2018; Li & Li, 2022). Allocating additional time to work for their course after school hours is unavoidable making this class more demanding than the traditional one.

The Unavailability of Immediate Guidance

As students work independently anywhere outside the classroom, the instructor encourages students to independently learn the content and complete the assigned exercises. Students need to work to comprehend the content by watching the provided videos, reading the slides, or web pages which is necessary to prepare them to get engaged in the upcoming classroom discussion. For this type of FCM, Cabi (2018) reported issues regarding the absence of immediate guidance. The author argued that the absence of direct support during students' self-learning made students feel frustrated when they encountered issues dealing with the content or technical matters.

DISCUSSION

This systematic review aimed to answer five research questions regarding the implementation of FCM. The first research question concerns the theoretical framework used to design FCM. The findings managed to identify several frameworks in the study such as constructivism which was explicitly integrated into the majority of experimental studies (e.g. Yu & Gao, 2022; Zhonggen, 2019; Cabi, 2018; Bonyadi, 2018; Yang et al., 2019; Vaezi et al). As for other studies adopting a case study design, various theories were applied to collect the data including inquiry-based learning in Bakla (2018) and the humanistic learning theory in Liu and Zhang (2018). One interesting point from the finding is that the application of FCM emphasized the types of technology used during the process and yet the strategies to optimize the classroom discussion lack attention in several studies. Thus, future research can address this issue and provide more comprehensive findings which covers both online and offline learning acticities.

The second question concerns the research methodologies and technological tools used to conduct studies of FCM. With an emphasis on empirical studies, this review found that a quasi-experimental design was employed in the majority of the study followed by the an experimental design, a case study, and classroom action research. No study adopted a phenomenological approach that might present valuable and in-depth insights regarding the application of FCM from

the perspectives of either instructors or students. The population was also dominated by undergraduate students with various majors from universities in Asia. It implies that there is still room for investigation regarding the level of students for research participants by recruiting master's students. Moreover, FCM utilized various tools such as online instructional videos, social media, mobile applications, and other instructional tools. However, studies with new technology such as augmented reality and artificial intelligence under the scheme of FCM might reveal interesting findings.

The third question attempted to describe the practical instructional benefits of FCM. The majority of the studies reported the effectiveness of the model in developing students' language skills (e.g. Abdullah et al., 2019; Bonyadi, 2018; Wu et al., 2017; Lin & Hwang, 2018) as well as their academic progress (e.g. Yang et al., 2019). They also indicated the adoption of this model on the improvement of students' motivation and learning efficacy by incorporating collaborative and self-directed learning. These benefits indicated that FCM emphasized the importance of students' independence to manage their own learning and their initiative to get actively engaged in learning discussion. As a consequence, teachers need to harness suitable strategies and approaches to provide necessary guidance and solution to any emerging issues.

The fourth question deals with the barriers to the implementation of FCM. Several studies reported the emerging instructional barriers that impeded the effective application of the model. For example, students encountered issues in managing their own learning activities prior to the class due to distractions from their surroundings and time management (Vuong et al., 2018; Cabi, 2018). It is made even more problematic with the unavailability of immediate guidance from instructors (Cabi, 2018). In terms of technical issues, the quality of videos or the internet connections proves to be significant barriers for students to comprehend the learning content (Vuong et al., 2018; Zhonggen, 2019; Teng, 2017; Cabi, 2018; Li & Li, 2022). Additionally, in some cases, both students and instructors were burdened with the additional workload resulting from the FCM application (Zhonggen, 2019; Suranakkharin, 2017; Bakla, 2018; Vuong et al., 2018; Li & Li, 2022). For teachers planning to adopt FCM, the ability to identify issues in the early stage of the application of this model is deemed necessary. Their efforts to identify and alleviate issues as early as possible will determine the success of this model to improve the quality of instruction.

Pedagogical Implications

It is reported in previous studies that implementing FCM can pose both instructors and students with various instructional challenges. For instance, the expected face-to-face discussion as the extension of the class might not run as expected because students do not comprehend the provided content and feel reluctant to consult their difficulties during the pre-class sessions (Cabi, 2018). Additionally, students might find it hard to manage their learning due to the absence of direct supervision from the instructor. For the instructor, managing students outside the classroom prove to be a difficult task as students have their respective personal matters. To gain success in this model, both instructors and students need to make a clear commitment to actively contribute to the actual instructional practices underlying this model. Instructors should also consider allowing students to interact with them outside the class in either personalized or collaborative interaction (Butt, 2014; Milman, 2012).

The instructors hold an essential role as learning facilitators who provide necessary feedback for learning and build students' positive perceptions of the learning process aside from preparing learning content. They should also guide students during their offline and online learning activities by helping them to construct a certain level of content comprehension during the process (Bonyadi, 2018). Akayoğlu (2021) highlighted the instructors' attitudes on the online platform which affect the overall application of FCM. Thus, they need to pay attention to students' characters to diminish feelings of anxiety or embarrassment amomng students in organizing the class activities. It is also necessary for them to formulate short assignments that include new information from students' learning content to guide them during their self-directed learning (Aprianto et al., 2020). Furthermore, Li and Li (2022) claimed that support from instructors and any involved parties in the education field is necessary to build students' positive attitudes toward FCM and thus help students to alleviate emerging issues during the application of this model.

The preparation stage of FCM is essential to get students accustomed to this approach. Suranakkharin (2017) and Liu and Zhang (2018) highlighted the need for the instructor to prepare students for the transition of instructional mode from conventional to FCM. They agreed that proper time allocation for the preparation stage is necessary for the transition. For this reason, a computer lab can be used to provide a tutorial for students to practice the accessibility of the applications or platforms used for the models (Lee & Wallace, 2018). Thus, instructors can create PowerPoint presentations or even distribute a hard copy of the materials for students to prevent possible technical issues dealing with the internet connection or the software. The online platform should be selected and utilized by considering the ease of access for various types of computing devices. Also, students should be directed to use the facility in the university library or public places that allow them to access their online content or class. Furthermore, special attention is deemed necessary for passive-type students to encourage them to participate in class discussions to obtain the instructional goals of FCM. In a similar vein, Aprianto et al. (2020) emphasized the need for instructors to allocate their time for students to consult their problems and obtain instant solutions that will decrease the risk of losing learning

interest during the class periods or outside the classes. Also, Liu and Zhang (2018) pointed out the importance of preparing students' mindsets for independent learning before proceeding to the application of this model. The model might be rendered ineffective due to wrong mindsets and the lack of preparation for the transition of pace. This claim is supported by Vaezi et al. (2019) who asserted that the success of FCM can be attributed to the fact that the time of learning has been extended by adding extra learning activities before the class to better prepare students with the instructional materials.

E-learning content needs to be carefully designed and developed to provide necessary information prior to the class meeting. As students need to work by themselves, the content should be organized in an attractive, clear, and brief fashion so students get attracted to watch and read the content. Instructors might need to creatively make a self-video recording or harness the available third-party videos from YouTube or other web pages that provide an explanation about the lesson and serve as a reference for the upcoming classroom discussion (Şengel, 2016). The videos containing the required information for flipped learning should be created as concisely as possible and preferably contain some engaging elements such as quizzes and games to attract students' attention without omitting its focus on content comprehension (Leis & Brown, 2018; Ekmekci, 2017). Moreover, it is necessary to consider the uploaded data sizes to ease the process of downloading or buffering. Alternatively, the instructors can store the file in a portable data storage (flash disk) and assign the class representative to handle the distribution of the file (Ekmekci, 2017). Considering the diversity of students' competence, Chen and Liu (2019) suggested that instructors consider designing an assignment that could be completed within one hour during the pre-class step. Both Li and Li (2022) and Yu and Gao (2022) pointed out the length of videos for learning content. More specifically, Yu and Gao (2022) agreed that the short videos (5 minutes or less) worked more effectively to improve students' learning outcomes, learning motivation, and participation in learning than longer videos. The lesson plan should cover learning activities not only inside but also outside the classroom.

Unlike the traditional instructional method, FCM heavily emphasizes students' role as the main actor in learning, thus encouraging them to get actively involved in classroom discussions. In some cases, this idea can be difficult. Therefore, Bonyadi (2018) argued that FCM should be incorporated as a complimentary instruction instead of a regular substitute for conventional in-person classes due to several technical restrictions that students might encounter such as internet connection and device requirements. Milman (2012) suggested incorporating FCM for class activities that focus on procedural knowledge (the knowledge to conduct a specific skill or task).

Research Implications

The direction of future studies can be directed to examine the application of FCM in different socio-cultural contexts. All studies derive from Asian regions such as Taiwan, China, Indonesia, and Turkiye, most of which consider English as a foreign language. Each country has its unique characteristics of learners and learning values. Conducting a collaborative study between two or several countries to compare the application of FCM might prove to be beneficial to identify the emerging issues in each country and generate valuable findings for future applications of this approach. Also, it is still difficult to find studies regarding FCM for ESL or EFL teaching that is originated in Africa, Europe, and The United States. More studies in different parts of the world that possess diverse socio-cultural backgrounds will be valuable references for future studies within this scope.

Other investigations can also address the adoption of mobile devices in the FCM setting. The ubiquity of mobile devices among students is an indication that mobility can be an essential element in learning. Various studies have provided evidence on how mobile devices can be harnessed to improve students' language skills and even some of which have combined the elements of FCM and mobile learning (e.g., Wu et al., 2017; Chen Hsieh et al., 2017). There are still various mobile learning platforms for the application of FCM such as WhatsApp, Instagram, and YouTube. Moreover, new technology such as augmented reality and artificial intelligence might provide new learning experiences for students during FCM. Future studies can be directed to investigate how these mobile platforms and new technology can affect students' language competence under the scheme of FCM.

Limitations

This study is expected to serve as a valuable reference for future studies regarding the adoption of FCM in the scope of EFL/ESL contexts in higher education. However, the findings of this study yield some limitations that can be improved in future studies. For instance, all reviewed articles originate from Asia especially East Asia. The issue of generalization can be a concern as the findings might not be applicable to other studies held in countries with different socio-cultural backgrounds.

Furthermore, the limited access to particular journals has been an essential issue in this study. The majority of the articles are retrieved from open-access journals with the inclusion of Scopus-indexed criteria to address the quality issue. There are still many high-quality journals that might offer valuable information for the scope of this study. Thus, future studies can include articles from these journals for a more detailed and extensive discussion.

CONCLUSIONS

FCM has shown its potential as an alternative instructional model that managed to elevate the quality of EFL/ESL learning. Moreover, multiple technological tools including teaching platforms and multimedia content were utilized to deliver learning content to learners. However, there is an indication of some technical and non-technical issues which impede the successful integration of this model. Future studies can address these issues to formulate strategies that provide solutions to these issues. It will also be interesting to describe the effects of mobile learning platforms and new technology such as augmented reality and artificial intelligence in the application of FCM.

From the previous research, this review formulates pedagogical implications for the future implementation of FCM. First, instructors need to initially identify students' familiarity with the medium used in FCM as they need to make sure that the model is practical for the class. Second, content should be made in a sort, concise, comprehensive, and attractive fashion so that students will be encouraged and need to make little effort to learn the content. Third, the lesson plan should be made creatively to engage students in an interactive discussion that occurs not only inside the classroom but also outside the classroom. To sum up, this review strengthens the idea of not lagging in the adoption of the latest instructional models and technological innovations. It also highlights the essence of students' and teachers' readiness for the latest trends in the educational field.

DECLARATION OF COMPETING INTEREST

None declared.

REFERENCES

Abdullah, M. Y., Hussin, S., & Ismail, K. (2019A). Implementation of flipped classroom model and its effectiveness on English speaking performance. *International Journal of Emerging Technologies in Learning*, *14*(9), 130–147. https://doi.org/10.3991/IJET.V14I09.10348

Abdullah, M. Y., Hussin, S., & Ismail, K. (2019B). Investigating the effects of the flipped classroom model on Omani EFL learners' motivation level in English speaking performance. *Education and Information Technologies*, *24*(5), 2975–2995. https://doi.org/10.1007/s10639-019-09911-5

Abdullah, M., Hussin, S., & Ismail, K. (2021). Does flipped classroom model affect EFL learners' anxiety in English speaking performance?. *International Journal of Emerging Technologies in Learning*, *16*(1), 94-108. https://doi.org/10.3991/ijet.v16i01.16955

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Education Research and Development, 34*,1–14. https://doi.org/ 10.1080/07294360.2014.934336

Akayoğlu, S. (2021). Teaching CALL to pre-service teachers of English in a flipped classroom. *Technology, Knowledge and Learning*, *26*(1), 155-171. https://doi.org/10.1007/s10758-019-09428-3

Aprianto, E., Purwati, O., & Anam, S. (2020). Multimedia-assisted learning in a flipped classroom: A case study of autonomous learning on EFL university students. *International Journal of Emerging Technologies in Learning*, *15*(24), 114. https://doi.org/10.3991/ijet.v15i24.14017

Artino, A. R., & Stephens, J. M. (2009). Academic motivation and self-regulation: A comparative analysis of undergraduate and graduate students learning online. *The Internet and Higher Education, 12*(3), 146–151.

Bakla, A. (2018). Learner-generated materials in a flipped pronunciation class: A sequential explanatory mixed-methods study. *Computers and Education*, *125*, 14–38. https://doi.org/10.1016/j.compedu.2018.05.017

Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education.

Bonyadi, A. (2018). The effects of flipped instruction on Iranian EFL students' oral interpretation performance. *Journal of Asia TEFL*, *15*(4), 1146–1155. https://doi.org/10.18823/asiatefl.2018.15.4.19.1146

Butt, A. (2014). Student views on the use of a flipped classroom approach: Evidence from Australia. *Business Education & Accreditation, 6*(1), 33e43.

Cabi, E. (2018). The impact of the flipped classroom model on students' academic achievement. *International Review of Research in Open and Distance Learning*, *19*(3), 202–221. https://doi.org/10.19173/irrodl.v19i3.3482

Chen Hsieh, J. S., Huang, Y. M., & Wu, W. C. V. (2017). Technological acceptance of LINE in flipped EFL oral training. *Computers in Human Behavior*, *70*, 178–190. https://doi.org/10.1016/j.chb.2016.12.066

Chen, S. L., & Liu, Y. T. (2019). High and low achievers' accepted workload, preferred work form, and perceived usefulness in flipped classrooms. *Journal of Asia TEFL*, *16*(1), 294–306. https://doi.org/10.18823/asiatefl.2019.16.1.19.294

Chen, F., Lui, A. M., Martinelli, S. M. (2017). A systematic review of the effectiveness of flipped classrooms in medical education. *Medical Education, 51*(6), 585-597. http://doi. org/10.1111/medu.13272

Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. University of Rochester Press.

Deng, F. (2019). Literature review of the flipped classroom. *Theory and Practice in Language Studies, 9*(10), 1350-1356. http://dx.doi.org/10.17507/tpls.0910.14

Doman, E. (2016). "Flip" or not to "Flip": That is the question. In E. Dorman & J. Bidal (Eds.), *Departing from tradition: Innovations in English language teaching and learning* (pp.45-61). Cambridge Scholars Publishing.

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education, 15*(1), 7-23.

Garrison, D. R., Cleveland-Innes, M., & Fung, T. S. (2010). Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. *The Internet and Higher Education, 13*, 31-36.

Gaughan, J. E. (2014). The flipped classroom in world history. *History Teacher, 47*(2), 221-244.

Halili, S. H., & Zainuddin, Z. (2015). Flipping the classroom: What we know and what we don't. *The Online Journal of Distance Education and e-Learning, 3*(1), 28–35.

Higgins, J. P. T., & Green, S. (Eds). (2008). *Cochrane handbook for systematic reviews of interventions: Cochrane book series*. Wiley-Blackwell.

Hung, H. T. (2017). Design-based research: Redesign of an English language course using a flipped classroom approach. *TESOL Quarterly*, *51*(1), 180–192. https://doi.org/10.1002/tesq.328

Jensen, B. A. (2019). Using flipped learning to facilitate cross-cultural critical thinking in the L2 classroom. *Die Unterrichtspraxis / Teaching German*, *52*(1), 50–68. https://doi.org/10.1111/tger.12084

Johnson, R.T., & Johnson, D.W. (1986). Cooperative learning in the science classroom. *Sci. Child, 24*, 31–32.

Karami, A. (2021). The Schematic Information-Processing (SIP) model of reading comprehension: Theoretical support for the utilization of text-relevant video segments to teach culturally unfamiliar texts in second/foreign language classrooms. *Cogent Education, 8*(1), 1-25. https://doi.org/10.1080/2331186X.2021.1891613

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *Journal of Economic Education, 31*(1), 30-43.

Lee, G., & Wallace, A. (2018). Flipped learning in the English as a foreign language classroom: Outcomes and Perceptions. *TESOL Quarterly*, *52*(1), 62–84. https://doi.org/10.1002/tesq.372

Leis. A., & Brown, K. (2018). Flipped learning in an EFL environment: Does the teacher's experience affect learning outcomes? *The EUROCALL Review, 26*(1), 3-13. https://doi.org/10.4995/eurocall.2018.8597

Li, Z., & Li, J. (2022). Using the flipped classroom to promote learner engagement for the sustainable development of language skills: A mixed-methods study. *Sustainability*, *14*(10), 1-16. https://doi.org/10.3390/su14105983

Li, Y., & Qu, C. (2019). College English education platform based on browser/server structure and flipped classroom. *International Journal of Emerging Technologies in Learning*, *14*(15), 171–181. https://doi.org/10.3991/ijet.v14i15.11147

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P., J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine, 62*, e1-e34.

Lin, C. J., & Hwang, G. J. (2018). A learning analytics approach to investigating factors affecting EFL students' oral performance in a flipped classroom. *Educational Technology and Society*, *21*(2), 205–219.

Lin, C.-J., Hwang, G.-J., Fu, Q.-K., & Chen, J.-F. (2018). A flipped contextual game-based learning approach to enhancing EFL students' English business writing performance and reflective behaviors. *Educational Technology & Society, 21*(3), 117–131. http://www.jstor.org/stable/26458512

Liu, X., & Zhang, J. (2018). Application of computer distance education in practical English writing teaching. *International Journal of Emerging Technologies in Learning*, *13*(4), 71–82. https://doi.org/10.3991/ijet.v13i04.8471

Milman, N. B. (2012). The flipped classroom strategy: What is it and how can it best be used? *Distance Learning, 9*(3), 85–87.

Öztürk, M., & Çakıroğlu, Ü. (2021). Flipped learning design in EFL classrooms: Implementing self-regulated learning strategies to develop language skills. *Smart Learning Environments*, *8*(1), 1-20. https://doi.org/10.1186/s40561-021-00146-x

Öztürk, M., & Ünal, Ç. (2021). Flipped learning design in EFL classrooms: Implementing self-regulated learning strategies to develop language skills. *Smart Learning Environments, 8*(2), 1-20. http://dx.doi.org/10.1186/s40561-021-00146-x

Powell, K. C., & Kalina, C. J. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education, 130*(2), 241-250.

Santikarn, B., & Wichadee, S. (2018). Flipping the classroom for English language learners: A study of learning performance and perceptions. *International Journal of Emerging Technologies in Learning*, *13*(9), 123–135. https://doi.org/10.3991/ijet.v13i09.7792

Şengel, E. (2016). To FLIP or not to FLIP: Comparative case study in higher education in Turkiye. *Computers in Human Behavior, 64*, 547–555. https://doi.org/10.1016/j.chb.2016.07.034

Shyr, W. J., & Chen, C. H. (2018). Designing a technology-enhanced flipped learning system to facilitate students' self-regulation and performance. *Journal of Computer Assisted Learning*, *34*(1), 53–62. https://doi.org/10.1111/jcal.12213

Suranakkharin, T. (2017). Using the flipped model to foster Thai learners' second language collocation knowledge. *3L: Language, Linguistics, Literature*, *23*(3), 1–20. https://doi.org/10.17576/3L-2017-2303-01

Teng, M., F. (2017). Flipping the classroom and tertiary level EFL students' academic performance and satisfaction. *Journal of Asia TEFL, 14*(4), 605-620. doi:http://dx.doi.org/10.18823/asiatefl.2017.14.4.2.605

Vaezi, R., Afghari, A., & Lotfi, A. (2019). Investigating listening comprehension through flipped classroom approach: Does authenticity matter? *CALL-EJ*, *20*(1), 178–208.

Vuong, N. H. A., Tan, C. K., & Lee, K. W. (2018). Students' perceived challenges of attending a flipped EFL classroom in Vietnam. *Theory and Practice in Language Studies*, *8*(11), 1504. https://doi.org/10.17507/tpls.0811.16

Wu, W. C. V., Chen Hsieh, J. S. C., & Yang, J. C. (2017). Creating an online learning community in a flipped classroom to enhance EFL learners' oral proficiency. *Educational Technology and Society*, *20*(2), 142–157.

Yang, Liu, & Todd. (2019). Effects of flipped classroom on high- and low-achievers' English vocabulary learning. *The Journal of Asia TEFL*, *16*(4), 1251–1267. http://dx.doi.org/10.18823/asiatefl.2019.16.4.12.1251

Yu, Z., & Gao, M. (2022). Effects of video length on a flipped English classroom. *SAGE Open*, *12*(1), 1-14. https://doi.org/10.1177/21582440211068474

Yu, Z., & Yu, L. (2017). Correlations between learners' initial EFL proficiency and variables of clicker-aided flipped EFL class. *Education and Information Technologies*, *22*(4), 1587–1603. https://doi.org/10.1007/s10639-016-9510-0

Zainuddin, Z., Habiburrahim, Muluk, S., & Keumala, C. M. (2019). How do students become self-directed learners in the EFL flipped-class pedagogy? A study in higher education. *Indonesian Journal of Applied Linguistics*, *8*(3), 678–690. https://doi.org/10.17509/ijal.v8i3.15270

Zhang, H. (2019). A flipped classroom-based education system for college English teaching. *International Journal of Emerging Technologies in Learning*, *14*(16), 120–129. https://doi.org/10.3991/ijet.v14i16.11152

Zhonggen, Y. (2019). Video lecturing in Clicker-assisted English flipped class. *PLoS ONE*, *14*(10), 1–18. https://doi.org/10.1371/journal.pone.0224209

Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 1–19). Guilford Press.

https://doi.org/10.17323/jle.2022.14433

Modelling the Linguistic Worldview: Subject Field Scoping Review

Anna Zakharova, Tatyana Baranovskaya, Marina Antonova, Stanislav Li

HSE University , Moscow, Russia

ABSTRACT

Background. The linguistic worldview theory stems from Humboldt's ideas concerning the interdependence of language and its speakers. Since Humboldt's time national linguistic worldviews remain a challenging aspect of linguistic research accounting for a significant number of publications. As linguistic worldview is a multi-faceted phenomenon, studies differ in terms of the facets they investigate and applied methodology. The linguistic worldview research has covered a wide array of linguistic worldview fragments and utilised monolingual material as well as cross-linguistic analysis of worldview fragments. However, so far, little attention has been paid to the analysis of this ever-increasing body of research and quite few studies have attempted to review the literature in this field.

Purpose. The subject field scoping review aims to overview the available literature on national linguistic worldviews and answers the three questions: What is a diachronic change of researchers' interest in the issue? What fragments of linguistic worldview have been investigated so far? What are the most frequently used methods of modelling the linguistic worldview?

Methods. The literature for the scoping review was retrieved from the three electronic databases: SciVerse Scopus, Web of Science (Core Collection), and Google Scholar. Literature selection was performed in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA). To generate a list of current research directions, which rest on the analysis of fragments of linguistic worldview, the visualisation map of authors' key words was created using the VOSviewer software.

Results and Implications. The concise scoping review of the previously published literature on linguistic worldview shows that most publications date between 2015 and 2021 and the years yielding most publications are 2018 and 2019. The number of publications mentioning linguistic worldview grew each year within the study period. The analysis of linguistic worldview fragments helped to identify six directions of researchers' interest in the linguistic worldview field which include lexical fragments, linguistic worldview and consciousness, phraseological fragments, reflection of linguistic worldview in different discourses, linguistic worldview in translation and teaching, and grammar fragments. The most frequently used methods of modelling the linguistic worldview are an associative psycholinguistic experiment, a conceptual analysis, and a comparative method.

KEYWORDS

fragments of linguistic worldview, language, linguistic worldview, modelling, scoping review

Citation: Zakharova A., Baranovskaya T., Antonova M., & Li S. (2022). Modelling the Linguistic Worldview: Subject Field Scoping Review. *Journal of Language and Education, 8*(3), 150-166. https://jle.hse.ru/article/view/14433

Correspondence:
Anna Zakharova,
annazakharova@hse.ru

Received: June 03, 2022
Accepted: September 17, 2022
Published: September 30, 2022

INTRODUCTION

The linguistic worldview theory is rooted in Humboldt's ideas concerning the interdependence of language and its speakers. While language reflects various aspects of a community life, thus accumulating the nation's spirit and mindset, it concurrently shapes human consciousness. Every language therefore represents its community unique view and interpretation of the outside world through its specific system. As Wilhelm von Humboldt claims, language is the outward appearance of the spirit of peoples, their language is their spirit, their spirit is their language (Humboldt, 1949, p. 41).

Understanding of the linguistic worldview concept and directions of its research are closely connected with its historical background. Humboldt's ideas of language content ethnicity came to the

linguists' spotlight in the 1930s, in particular, Leo Weisgerber developed these ideas and put forward the theory of the linguistic worldview (*sprachliches Weltbild*). His first definition of this concept was word-centric for he argued that the vocabulary of a particular language comprises besides linguistic signs a set of conceptual thinking tools which native speakers master while studying this vocabulary (Weisgerber, 1930). In the article *Sprache*, Leo Weisgerber endeavors to clarify the concept of linguistic worldview, emphasising the nation's spiritual content that, being the treasure of knowledge, 'lives' in the language of a particular community (Weisgerber, 1931). Every language incorporates the totality of formalised knowledge, which the members of the linguistic community have developed since the existence of this particular linguistic community on the basis of their linguistic ability and presented in linguistic forms (Weisgerber, 1931). Neo-Humboldtians aimed at unearthing the cognitive content embedded in a given language. They claimed that language should be viewed not only as a means of communication but as a creative power of the spirit, as a "symbolic guide to culture" (Sapir, 1929, p. 70), as a means of accessing the world (Christmann, 1967).

'Accessing' the world is governed by conceptualisation and categorisation, which are culturally dependent and vary across languages (Lakoff, 1987; Lakoff & Johnson, 2003). When verbalising concepts, an appropriate method of linguistic representation is chosen, which implies categorisation of the designated, i.e., attributing it to the class of animate or inanimate objects, actions or states, qualities and so forth, so that the designation takes the form of one of the available word classes. Therefore, the concept of linguistic worldview implies "a language-entrenched interpretation of reality… The interpretation is a result of subjective perception and conceptualization of reality performed by the speakers of a given language; thus, it is clearly subjective and anthropocentric but also intersubjective (social)" (Bartmiński, 2009, p. 23). A set of ideas about the world is formed historically and inherited by subsequent generations.

National linguistic worldviews have been actively investigated since about the 1980s (Berghout, 2006). The term *sprachliches Weltbild*, literally *linguistic "picture" of the world*, has its equivalents in other languages: *linguistic worldview* in English, *jazykový obraz světa* in Czech, *językowy obraz świata* in Polish, *yezichka slika sveta* in Serbian, *movna kartyna svitu* in Ukranian, or *yazykovaya kartina mira* in Russian (Glaz, 2022). The languages differ in the number of word classes, modes of verbalising the same concepts, language represents "a segmentation of the world which is not inherent in things but precisely in language" (Christmann, 1967). This accounts for the directions of research which has covered a wide array of linguistic worldview fragments, such as various lexico-semantic groups, idioms representing certain concepts; grammar forms and syntactic structures that govern the expression of the idea; ways of categorising things and phenomena as well as ways of conceptualising abstract notions (Mierzwińska-Hajnos, 2010; Daulet et al., 2019; Bolshakova & Gladkova, 2020; Humeniuk, 2021; Litvishko et al. 2022). Studies have utilized either monolingual material or involved cross-linguistic analysis of worldview fragments (Barczewska et al., 2016; Alyoshin, 2018; Rudenka & Fang, 2018; Tóth, 2018; etc.).

Providing insight into national linguistic worldviews still remains a challenging aspect of linguistic research. Since linguistic worldview is a multi-faceted phenomenon, studies differ in terms of the facets they investigate and applied methodology. However, so far, little attention has been paid to the analysis of this ever-increasing body of research. To the best of our knowledge, quite few studies have attempted to review the literature. Sartini and Ahimsa-Putra (2017) focused on the studies analysing various worldviews, with little regard to language. The reviewed studies were grouped according to the research paradigms they followed, namely structuralism, evolutionism, functionalism, ethno-science, cultural interpretation and so forth. Haydruk (2018) briefly covered general ideas about linguistic worldview suggested by such prominent scholars as Osgood, Wierzbicka, Palmer and Budagov. So, modern theoretical and empirical studies as well as their quantitative and release date aspects remained beyond the scope of these reviews.

This paper seeks to fill this gap by overviewing available literature on national linguistic worldviews. This study's prior objectives are (1) to map the existing literature and (2) provide a succinct scoping review of the linguistic worldview domain. Following these objectives, the three research questions are put forward:

(1) What is a diachronic change of researchers' interest in the issue?
(2) What fragments of linguistic worldview have been investigated comprehensively?
(3) What are the most frequently used methods of modelling the linguistic world view?

The main contribution of this study is therefore twofold: the systematic analysis of the papers investigating national worldviews is presented; this analysis is based not on subjective selection of data but on that arranged according to the PRISMA statement (Preferred Reporting Items for Systematic reviews and Meta-Analyses). The proposed scoping review is likely to stimulate researchers for theoretical and empirical analysis of those linguistic worldview fragments that still remain underinvestigated. Moreover, this may advance further systematic reviews of the linguistic worldview domain.

METHODS

This scoping review was undertaken using the framework proposed by Arksey and O'Malley (2005), which includes five

phases: defining the research questions, selecting relevant literature, specifying criteria for inclusion and exclusion to make a sample, recording the results across sources, and summarizing the findings. It was found that some researchers expressed concerns about Arksey and O'Malley's framework's limitation to provide for an assessment of the quality of the selected literature (Peters et al., 2015; Panic, 2013; Moher et al., 2009). Given the significance of quality assessment for scoping studies, the authors followed the recommendations provided by Levac et al. (2010) and Daudt et al. (2013). Specifically, we applied carefully selected criteria for including/excluding sources for the present scoping review, used the PRISMA-ScR statement to ensure methodological and reporting quality of this scoping review and VOSviewer software to construct and visualise co-occurrence networks of key words extracted from the selected literature.

Search Strategy

To identify sources for inclusion, three electronic databases were searched: SciVerse Scopus, Web of Science (Core Collection), and Google Scholar. The publication years were limited to the period 2004 to the present because in Scopus and Web of Science, works mentioning linguistic worldview started to appear since 2004. The earlier papers considering linguistic worldview cannot be tracked in these two databases due to their years coverage, as they register papers published since 1996 and 1990 respectively. In Google Scholar the earliest works devoted to linguistic worldview dates back to 2005. However, the national linguistic worldviews have been in the researchers' focus since the 1980s (Berghout, 2006, p. 118).

The search focused on articles in peer-reviewed journals, conference proceedings, as well as books/book chapters that discuss the linguistic worldview. This scoping review covered works from all geographic locations. For feasibility purposes, we sought to identify literature published in English; all non-English content was excluded because of the time and cost needed for their translation. No grey literature was incorporated in our analysis. Moreover, if the analysis of the paper showed that the construct of linguistic worldview was used as one of the dependent variables without studying it, such paper was excluded from reviewing.

The terms used in the search were "linguistic worldview" and "linguistic world image". Mesh search for these terms in the aforementioned databases showed that these terms were mainly adopted within the social sciences including linguistics, psycholinguistics and cultural studies. As our research interest focuses on the concept of "worldview" as a phenomenon characteristic of linguistics, the key words were selected as entries in searching for the relevant sources to exclude studies irrelevant to the topic. The key word combinations included: "linguistic worldview" AND "linguistic world image". All the works in which one of these word combinations is present either in a title or abstract were considered linguistic-worldview-related literature.

Screening and Data Extraction

First, a total of 1,403 papers were found in the three databases. Duplicates (n = 1,130) were removed, and the remaining papers (n = 273) were analysed. Application of exclusion criteria resulted in exclusion of 105 records presenting non-English content. The iterative search resulted in 168 studies.

Second, after screening full-texts (n = 168), 45 records using linguistic worldview as a dependent variable were excluded as well as 7 records that failed to present methods applied by an author(s). This resulted in 116 sources that were subjects for data extraction.

For the selected articles (n = 116) we extracted author's name, author's country, year of publication, study design, and journal (source) name, text evidence describing a fragment of linguistic worldview, and methods applied by researchers. Then, the extracted information was exported to Excel for further analysis.

To illustrate the process of literature selection, the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) flow diagram was used (Figure 1).

Data Analysis

The data analysis was conducted in the following areas: fragments of linguistic worldview such as lexico-semantic word groups, phraseological units, grammar forms, syntactic structures, etc. to identify the areas of research in the field and methods applied by researchers, and the diachronic change of researchers' interest in linguistic worldview. Both quantitative (frequencies and percentages) and qualitative (text inspection) methods were applied for data analysis.

To generate a list of potential areas of research, firstly, the visualisation map of authors' key words was created with VOSviewer software (Figure 2); secondly, we scanned the titles and abstracts of exported articles.

The visualisation map of authors' keywords showed that the most frequently encountered author keywords were semantics, language, linguistics, metaphor, discourse, concept, human, translation. Then, after scanning the titles and abstracts we could more specifically define the areas of research: lexico-semantic groups; idioms representing certain concepts; grammar forms and syntactic structures; ways of categorizing things and phenomena; ways of conceptualizing abstract notions; discourse representation; translation; foreign language teaching.

Figure 1

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Flow Diagram Summarising Literature Selection

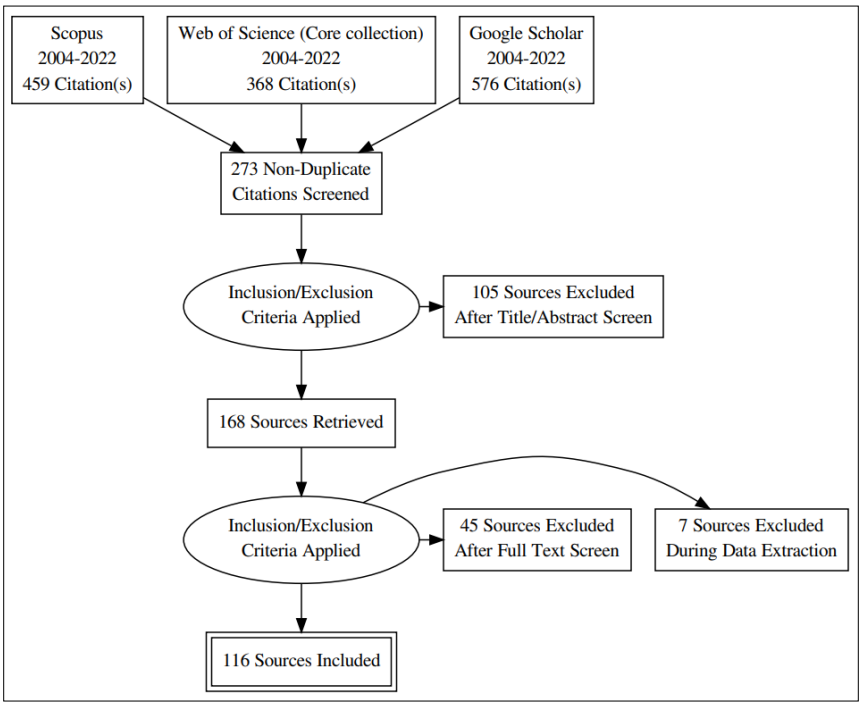
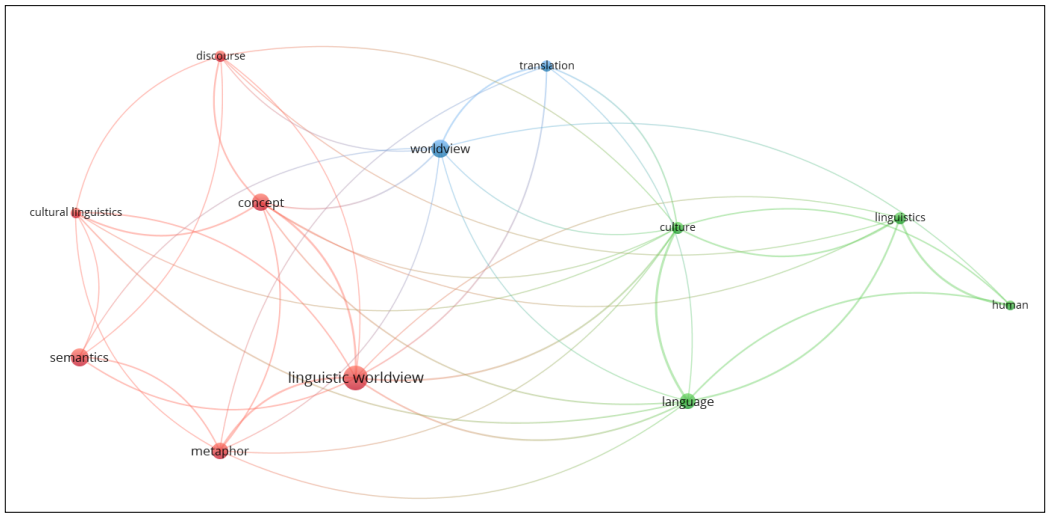


Figure 2

Network visualisation map of authors' keywords



To analyse the methods used by researchers and detect the diachronic change of scientists' interest we charted the main informational fragments. To do this we extracted the data from their original context and sorted materials according to the key themes. Then, the charted data was transferred to an Excel matrix. To categorise the context of the selected studies, categories of criteria were chosen: bibliographical data, direction of research, and methods used (Table 1).

The final stage of the present study includes classifying, summarising and disseminating the results of the data analysis. Presentation of findings obtained from the charted data includes: first, the description of bibliographical nature (author, author's country) of the selected studies and the research directions; second, methodological character of these studies (study design, methods applied by researchers to explore linguistic worldview, and data used for their analysis); third, the description of the diachronic change of scientists' interest in exploring linguistic worldview for the research time period (year of publication). This presentation of findings provides bases for detecting the key areas of research interest on linguistic worldview. The identified themes are reported in detail in the following sections.

Table 1

Categories of Variables with Their Subcategories Used in Coding of the Selected Literature

Bibliographical data	Directions of research	Methods used
Author(s)' name(s)	lexico-semantic group	Methods used: quantitative, qualitative or theoretical
Year of publication	idioms representing certain concepts	Timeframe for data collection
Title of the article	grammar forms and syntactic structures	Study design
Name of the journal	ways of categorising things and phenomena	
	ways of conceptualising abstract notions	
	discourse representation	
	translation	
	foreign language teaching	

RESULTS

Review of the Linguistic Worldview Research Domain

This section presents classification of the selected studies on linguistic worldview. Our analysis reveals the growth of research field over time, compares the ratios of fundamental and empirical research, enlists investigated fragments of linguistic worldview that illustrate the directions of research and finds out the methods applied by researchers.

Growth of Research Field over Time

Table 2 shows the diachronic change of researchers' interest in linguistic worldview issue over the period 2004-2022. Approximately, 83.6% (97) of the included articles were published in the last 7 years (2015-2021) of the study period. The number of publications mentioning linguistic worldview increased each year from 2014 to 2021. The years of 2018-2019 saw the largest number of works: 19 and 21 papers respectively. For illustration purposes, the years yielding no publications were excluded from Table 2. Researchers from 9 countries were observed in the retrieved documents. The two countries with the highest research output in linguistic worldview are Russia and Poland. Russia ranked first (77; 66.4%) followed by Poland (25; 21.5 %). Figure 3 demonstrates that empirical studies significantly prevail over theoretical ones (101 vs 15 papers).

Papers were published in a quite wide array of journals founded in European, Asian, and North and South American countries, such as Armenia, Belgium, Canada, China, the Czech Republic, Germany, India, the Netherlands, Poland, Russia, Slovakia, and Venezuela.

Mostly, papers related to linguistic worldview were published in journals released in Russia, Poland, and China. (Table 3).

Findings on Linguistic Worldview

As a result of the review of works devoted to the description of various national linguistic worldview, it is possible to present the main directions of research.

The vast majority of studies (71%) are devoted to the description of one national linguistic worldview, focusing on Hebrew, English, Chinese, Polish, Russian, Tatar, Teleutian, French and Yakut languages. 25.7% of the studies conduct a comparative analysis of two national linguistic worldviews, the analysed languages being English/Polish, English/Russian, Chinese/Russian, Polish/Ukrainian, Slovak/Hungarian, Tatar/English, Swedish/Russian. Comparison of three linguistic worldviews is presented by a rather restricted number of works – 2.9%, the analysed languages being English/Chinese/Russian, Lao/Hungarian/Russian.

As for research directions, it is possible to distinguish six of them: lexical fragments, linguistic worldview and consciousness, phraseological fragments, reflection of linguistic worldview in discourses, linguistic worldview in translation and foreign language teaching, and grammar fragments. The overwhelming majority of empirical studies examine national linguistic worldviews represented via lexical units or via their connection with consciousness – 31.7% and 21.5% respectively. There are also works that deal with specific linguistic worldview fragments, in particular texts of the Bible, advertising and Internet communication.

Theoretical Findings on Linguistic Worldview

The main challenges in linguistic worldview research have been described by Bartmiński (2012, 2015). The researcher suggests that a national linguistic worldview is studied with the help of two procedures: semasiological, which comes from a word meaning to the denoted object and onomasiological, which comes from the object to its denotation. Bartmiński explains both procedures in detail, while discussing

Table 2

Characteristics of Publications Included in this Scoping Review

	Publications included into this scoping review (n = 116)	
Author(s)' country	Russia: 77	Spain: 1
	Poland: 25	Estonia: 1
	China: 4	Armenia: 1
	Slovakia: 3	Denmark: 1
	Ukraine: 3	
Study design	Empirical studies: 101	Fundamental studies: 15
	Year of publication	Empirical Studies	Fundamental Studies
	2004	1	0
	2009	0	1
	2010	3	0
	2013	2	3
	2014	8	0
	2015	10	4
	2016	6	2
	2017	7	0
	2018	18	1
	2019	20	1
	2020	13	2
	2021	12	0
	2022	1	1

Figure 3

The Number of Fundamental and Empirical Publications per Year.

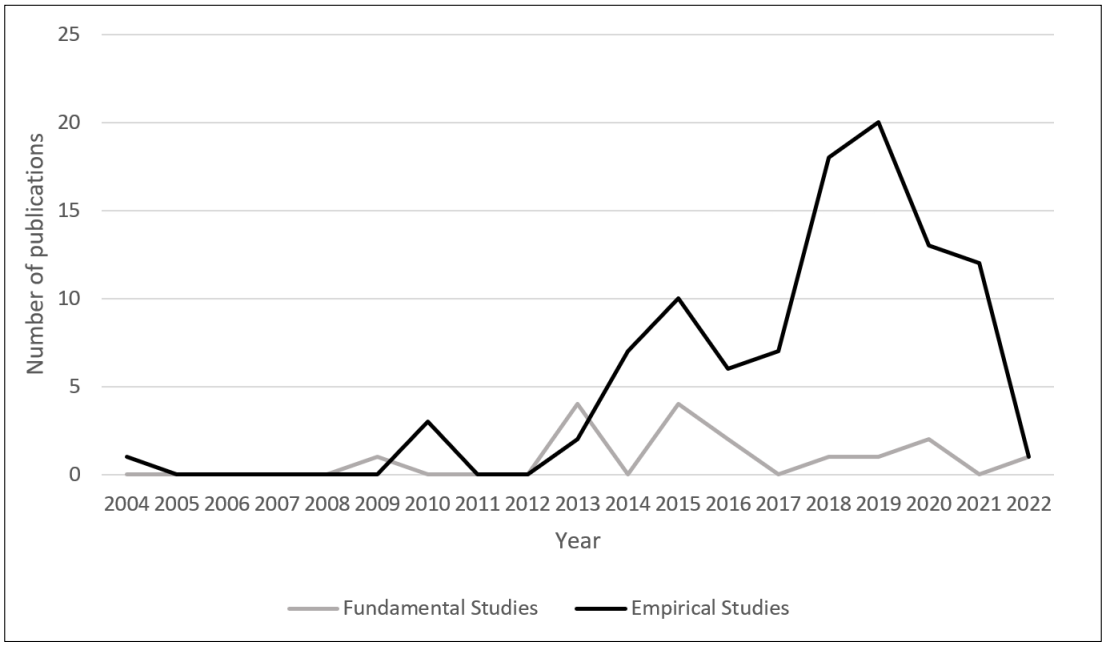


Table 3

Country of Publishing a Journal

Country	Journal
Russia	Voprosy Jazykoznanija
	Voprosy Kognitivnoy Lingvistiki
	Vestnik of Saint Petersburg University. Language and Literature
	Vestnik of Northern (Arctic) Federal University
	Vestnik Tomskogo Gosudarstvennogo Universiteta
	RUDN Journal of Language Studies, Semiotics and Semantics
	Journal of Siberian Federal University. Humanities & Social Sciences
	Sibirskii Filologicheskii Zhurnal
	Modern Journal of Language Teaching Methods
	Psycholinguistics
Poland	Acta Baltico-Slavica
	Cognitive Studies | Études cognitives
	Poradnik Jezykowy
	Półrocznik Językoznawczy Tertium
	Poznań Studies in Contemporary Linguistics
	Studia z Filologii Polskiej i Słowiańskiej
China	Journal of Suzhou University of Science and Technology
	Journal of University of Electronic Science and Technology of China
	Chinese Science and Technology Translators Journal

the type of data that should secure the validity of findings and the types of linguistic worldview fragments.

The book (Głaz et al., 2013) comprises studies concerning diverse aspects of linguistic worldview, in particular, presenting linguistic worldview in literature, stereotypes and values in a national linguistic worldview, the onomaseologiacal category of quantity, representation of such concepts as 'slave' and 'patriotism'.

The fundamental book by Głaz (2022) outlines the concept of linguistic worldview along with its historical background, dating back its origin to ancient Greece and tracing its development onwards in Western ethnosemantics and the theory of linguistic relativity. It surveys and compares the existing approaches to linguistic worldview analysis, the most propagating being universalism and language-specificity, cultural linguistics methods, cognitive ethnolinguistics methods, the Russian approaches to 'yazykovaya kartina mira'. The book pinpoints their differences convergence, and discrepancies. As a sample, the book proposes studies of the following three cases: Earth interpreted as home, mother and sister in speeches of Pope Francis; the family metaphor and the house metaphor.

Also, researchers differentiated between conceptual and linguistic worldviews (Grzegorczykowa, 2015; Ajdarkhan, 2018; Tokarski, 2016; Novoseletska et al., 2020). On the basis of theoretical studies, the concept of linguistic worldview is clarified, the dynamic of its development as well as its main universal features are described (Gabbasova et al., 2013; Suleimenova, 2013; Alefirenko & Li, 2015; Kozlova, 2016; Alefirenko et al., 2019). Apart from that, it was proposed to apply a questionnaire method to study a national linguistic worldview (Stria, 2015). According to the Polish ethnolinguistic school of Lublin, both language and cognition are embedded in culture, which accounts for formation of stereotypes that can be modelled via image schemas (Carisio, 2020).

Empirical Findings on Linguistic Worldview

Linguistic Worldview and Consciousness. Empirical findings in this section are presented according to the topical principle.

The correlation between linguistic worldview and human consciousness has been studied via concept analysis (11 papers), an associative experiment (6 papers) or comparative analysis of two national linguistic worldviews (6 papers). These studies concerned the Russian language except for only one research that focused on the Lao language (Lenart & Markovina, 2021). The experiments were carried out with a different number of respondents from 4 to 65 years old.

Depending on the age of the respondents, different formats of tasks were offered: filling out a paper or electronic questionnaire.

The interview of preschool age Lao speakers was conducted via the shoulder-to-shoulder method. The obtained data was compared with the data obtained during a previous study of Hungarian and Russian that used the same method and the same stimulus words. The data was downloaded and analysed with the help of the Sketch Engine (a tool for corpus research of language data). The associations received from the respondents, on the one hand, prove the similarity of the linguistic worldview of representatives of the three countries, on the other hand, demonstrate the unique characteristics of the national linguistic consciousness of Lao, Hungarian and Russian children (Lenart & Markovina, 2021). Also, through interviews of 4–6-year-old children, categorisation of artifacts in the Russian linguistic worldview was studied on the linguocognitive category TOYS (Dziuba & Podsukhina, 2019). The paper (Salnikova & Asanbaeva, 2017) presents linguistic representation of the associative area of the word "lyceum" in the linguistic worldview of a child living at the beginning of the XXI century. Another experiment addressing stereotypes about Great Britain and France in the linguistic worldview of philology students from Russia revealed 23 stereotype semantic fields (Lasitsa, 2020). Associative areas of the words denoting domestic animals were determined for the Russian and Mongolian linguistic worldviews (Shkuropatskaya & Undarmaa, 2016). The associative field with its core and periphery was identified for the lexeme 'client' (Kushmar, 2019).

Linguistic worldviews can intersect or integrate in the case of a bilingual speaker (Khayrullina et al., 2016) or closely related nations (Pivovar, 2019). The distinguishing features of FAMILY was compared for the Russian and Chinese languages (Shulgina, & Fang, 2014). Also, researchers examined the representation of MACHINE in English and Russian (Bulgakova & Sedelnikova, 2015) and FEMALE in English and Russian Orthodox Discourse (Baimuratova, & Korobeynikova, 2020). The concept 'female' was analysed across three languages as well – English, Russian and Tartar (Bazarova, & Gilyazeva, 2018).

As for concept analysis, representation, characteristic features and axiological value of the following concepts in national linguistic worldviews were analysed:

- concepts related to the idea of the anarchic denial in Russian (Martynov, 2014);
- HOSPITALITY in Kazakh (Chakyroglu & Suiyerkul, 2014);
- WEDDING in Russian (Komarova et al., 2017);
- concepts related to the conceptual sphere "geopolitics" in English (Kurbatova, 2017).
- HOLY FOOL in the linguistic worldview of the Russian-speaking population of Ukraine (Gordiienko-Mytrofanova & Kobzieva, 2018);
- HOME, FAMILY, WORK, EUROPE, FREEDOM and HONOUR in Lithuanian (Sawaniewska-Mochowa, 2018);
- LIE in French and Russian (Shatilova, 2018);
- FACE in Chinese (Daulet, 2019);
- LIGHT in Russian Medieval Religious Consciousness (Erofeeva, et al., 2019);
- MONEY in French (Shapovalova, 2019);
- BOUNDARY presented in Russian song lyrics (Milovanova & Matrusova, 2020).

Lexical Fragments. Empirical findings included into this section are presented according to the topical principle.

Papers dealing with lexis total 28. Researchers selected the data and material from lexicographic sources and various texts (Mierzwińska-Hajnos, 2010; Kostina et al., 2014; Dalabayeva, 2014; Jakubowicz, 2015; Bolshakova & Gladkova, 2020, etc.) or during field expeditions to places of compact residence of native speakers (Obraztsova et al., 2019; Abdullayeva et al., 2019; Olenev et al., 2019).

A challenging aspect of linguistic worldview analysis appears to be the formation of word metaphorical meaning since it inspired quite a number of studies. Researchers' interest might be justified due to the fact that metaphors allow for revealing linguoculturally driven ways to interpret abstract concepts and find similarities between things, which help shed light on peculiarities of a linguistic community thinking. From theoretical perspective, metaphors were described as units of conceptual and linguistic worldviews (Sadykova & Kayumova, 2014); two different approaches to metaphor analyses were compared – the Ethnolinguistic School of Lublin and Conceptual Metaphor Theory (Zinken, 2004).

Empirical studies of metaphors dwell upon various lexico-semantic groups relating to various spheres of life; metaphorical symbols are embedded in the language and consciousness of a language community speakers. For example, research into Russian and Galician zoomorphic metaphors denoting livestock found out associative links between animals' activity and way of life and people's behaviour and style of life; the study showed that both languages utilise universal axiological stereotypes stemmed in mythology and biblical texts while there are few nationally restricted metaphorical associations (Merzlikina, 2021). Similarly, the experimental study of Russian and Mongolian words denoting pets distinguished cultural peculiarities that account for differences in linguistic worldviews (Davaa, 2015). Mentality, cultural values and national character appeared to be encoded by somas, which was demonstrated by the analysis of Chinese linguistic worldview (Daulet et al., 2019). In addition, groups of professionals can exploit terms that were coined due to metaphorization as is the case with many of

the English theatre terms that have originated from poetic metaphors (Churilova et al, 2020) or vintners' terminology (Zawisławska, 2015).

Apart from that, researchers have been actively investigating more specific fragments of national linguistic worldviews, such as animal names in Russian (Krylova, 2019) or plant names in the Polish language (Mierzwińska-Hajnos, 2010, 2013); names of mental disorders in English, Chinese and Russian (Rudenka & Fang, 2018); zoonyms in the Russian and Mongolian linguistic worldviews (Shkuropaskaya & Davaa, 2018); names of emotions in Russian and English (Olomskaya et al., 2018); names of intellectual abilities in Estonian (Degel, 2010); hospitality words and phrases in Russian (Chirich & Shtukareva, 2019); tombstone inscriptions for dogs in Polish (Urszula, 2016). Investigation of Swedish and Russian words denoting human movements found out that the Russian linguistic worldview maps this fragment in much more detail than the Swedish one (Alyoshin, 2018). In addition, it appeared that the noun and verb semantics in Russian implicitly include a certain fragment of linguistic worldview of the cultural community (Kolmogorova, 2015).

Also, within mainstream cultures researchers address dialectal worldviews, in particular, the lexical system of Orenburg dialects (Russia) reveals a significant number of Turkic borrowings associated with the nature of the area (Bekasova, 2019). Modeling the fragment "Fire" of Pskov dialects allowed identifying more than forty toponyms related to the concept of conflagration (Bolshakova & Gladkova, 2020). Besides, researchers investigated the Teleut that is spoken by an indigenous minority of Russia, the language community totals about 2,600 people so the language is on the verge of extinction. Several fragments of the Teleut linguistic worldview were described, namely lexemes that verbalize culturally significant frames – food, hunting, horse-breeding, family, labour, and life (Araeva et al., 2015; Obraztsova et al., 2019; Abdullayeva et al., 2019; Olenev et al., 2019).

Assimilation of loan words implies their integration in a foreign linguistic worldview, even biblical borrowings undergo some modification entailed by incorporation of biblical stories in a given culture, in particular, Bartmiński (2019) distinguishes five level of assimilation of the words borrowed in Polish.

Phraseological Fragments. Empirical findings in this section are based on the topical principle.

National worldview is also reflected and fixed in idioms and various paroemiological units – proverbs, sayings, riddles, slogans, wellerisms and so forth (15 papers). All scholars agree that the semantics of a phraseological unit correlates with the culture code shared by the speaker and listener. For instance, the analysis of English gender-marked phraseological units revealed that gender stereotypes govern the national linguistic worldview formation to a great extent encoding the inequality between men and women (Humeniuk, 2021). The comparative study of the gender representation in Tatar and English phraseological units revealed that the female component in both languages frequently conveys a negative connotation (Khuzina & Mukhtarova, 2018). Comparison of Russian and Swedish phraseological units with the colour component showed that most of them characterise human beings, with less than 20% possessing positive connotation (Konovalova & Basova, 2021). In fact, the overwhelming majority of phraseological units possess the axiological component (Markelova et al., 2016; Kotova & Raina, 2020). Studies of phraseological units helped determine cross-language similarities of linguistic worldviews in French and Russian (Makarova et al., 2018; Gasymova et.al., 2022). Also, it was shown that the emotive and sensonymic components of Tatar idioms is determined by Tatar history, culture and traditions (Khasanzyanova, 2018; Yahin et al., 2017; Sibgaeva et al., 2021); the same seems to be true for the Uzbek linguistic worldview (Yuldashev, 2020). According to the study of phraseological units with a numerical component in the Yakut language, these units describe the appearance of the heaven hero and the underground creature from the Olonkho (Zakharov et al, 2021). In Chinese, phraseological units with zoonyms reflect the roles of certain animals in the life of the Chinese ethnos (Daulet et al., 2019).

The study of Polish proverbs selected from a 4-volume academic collection demonstrated that the triad family-work-religion is considered as priorities (Gieroń-Czepczor, 2020). However, the researcher suggests that the role of proverbs in the formation of these values is declining since most of the analysed units concern the irrelevant phenomena and relationships, thus becoming outdated. The use of transformed proverbs in memes and advertising demonstrates the non-acceptance of existing restrictive norms. Apart from that, phraseological units representing the concepts WORK and LAZINESS in the Archi language were examined (Samedov & Gasanova, 2017). Also, researchers endeavoured to reconstruct the linguistic worldview of the past epochs, in particular, that of the Poles of the XVIII century (Kuryłowicz & Szamryk, 2021).

The experience of creating the phraseological Polish-Ukrainian dictionary was presented by Sosnowski & Tymoshuk (2017). The dictionary comprises modern phraseological units that reflect the current linguistic worldview of Poles and Ukrainians.

Grammar Fragments. Grammatical categories seem to reveal a connection with human cognitive activity, which was proved by the research results (8 papers). In particular, these studies examine the role of the grammatical category of animacy-inanimacy of nouns in creating the Russian linguistic worldview (Narushevich & Bak, 2021), gender markers in language (Gulik, 2015; Humeniuk, 2021; Khuzina &

Mukhtarova, 2018) and the function of the Russian pronoun 'I' in phraseological units (Graneva, 2021) as well as the opposition of the Russian pronouns 'we' and 'they' (Revenko & Osetrova, 2019). The medieval linguistic worldview was analysed in terms of its influence on the morphological system of verb tenses in the modern Slavic languages (Shumilo, 2021). Comparative analysis of morphology of two genetically and typologically different languages – Slovak and Hungarian – revealed that a long-term cultural convergence leads to cognitive analogy even in typologically different languages (Tóth, 2018).

Reflection of Linguistic Worldview in Discourses. Since any author is a native speaker of a certain language and belongs to a certain culture, s/he possesses a certain linguistic worldview that is inevitably reflected in created texts. That is why text analysis, in particular the analysis of fiction contributes to research of national linguistic worldviews (15 papers). From this perspective, Polish and Canadian literature were analysed (Pajdzińska, 2013; Ustinova, 2014). Linguistic worldview was studied on the basis of works of such eminent men of letters as Ch. Dickens (Pospelova et al., 2021), the Russian writers I. Bunin (Antipina, 2020), I. Ilf and E. Petrov (Fefelova et al., 2018), K. Aksakov (Kalashnikova, 2018), modern American writers of Chinese origin E. Tang and M. H. Kingston (Korovina et al., 2020).

The religious discourse reflects linguistic worldview as well, for example the "Hagiography of Protopope Avvakum" was analysed in terms of axiological representation of a linguistic personality (Mirzoeva, 2019). To reconstruct the biblical vision of man and woman rooted in the religious value system, researchers addressed the prophetic books of the Old Testament in Hebrew (Szarlej, 2020).

In turn, scientific texts experience the impact of national linguistic worldviews, which was shown in the study by Tagirova (2019). Analyzing a collection of English research papers' abstracts borrowed from different journals, the author determines the influence of culture on pragmatics of the texts.

It is a well-known fact that Internet platforms have become an unprecedentedly popular means of communication for people from all walks of life. That is why blogs, forums, and social networks clearly present national linguistic worldviews. For instance, the analysis of Polish and American teenagers' vocabulary demonstrated that teenagers conceptualise their ideas about school in different ways, which is due to the differences in their linguistic worldviews (Barczewska, et al., 2016).

There is little doubt that advertisement is affected by linguistic worldview, in particular, comparison of Chinese and Russian advertising slogans allowed for discriminating common and national-specific characteristics of advertising texts (Li et al., 2018).

In terms of newspaper texts, analysis of Polish and English articles reporting on the same event – the coming out of a Polish priest – also reveals features peculiar to each of the national linguistic worldviews (Wyrwa, 2017). Besides, English newspapers' headlines (Yergaliyev et al., 2015) and articles covering the issue of migration in the European Union (Sipko, 2018) were analysed in terms of linguistic worldview reflection.

Linguistic Worldview in Translation and Foreign Language Teaching. Since discourses reflect national linguistic worldview, literature being no exception, translators frequently face a serious problem of preserving and transferring linguistically driven peculiarities of a translated text (8 papers). From this perspective, Humboldt's ideas about language and translation were discussed (Qin & Zeng, 2010; Gu, 2016; Gao & Hua, 2021). Researchers focus on difficulties of translating poetry, in particular Gicala analyses English translations of Wisława Szymborska poems and presents a translation strategy, which demonstates the reconstruction of the linguistic worldview for the target culture (Gicala, 2013; 2021). Another translation problem concerns humour as it varies in cultures and is expressed through different language means, for example, untranslatable humour metaphors are examined in Danish (Levisen, 2019). The literary technique of estrangement is viewed as a means of translation, complementary to translation strategies of domestication and foreignisation. This technique was applied in the Russian translation of the works of modern Israeli writers (Valkova, 2018). Besides, researchers dwelt upon translation of science fiction from Chinese to English (Xiu & Jiang, 2018).

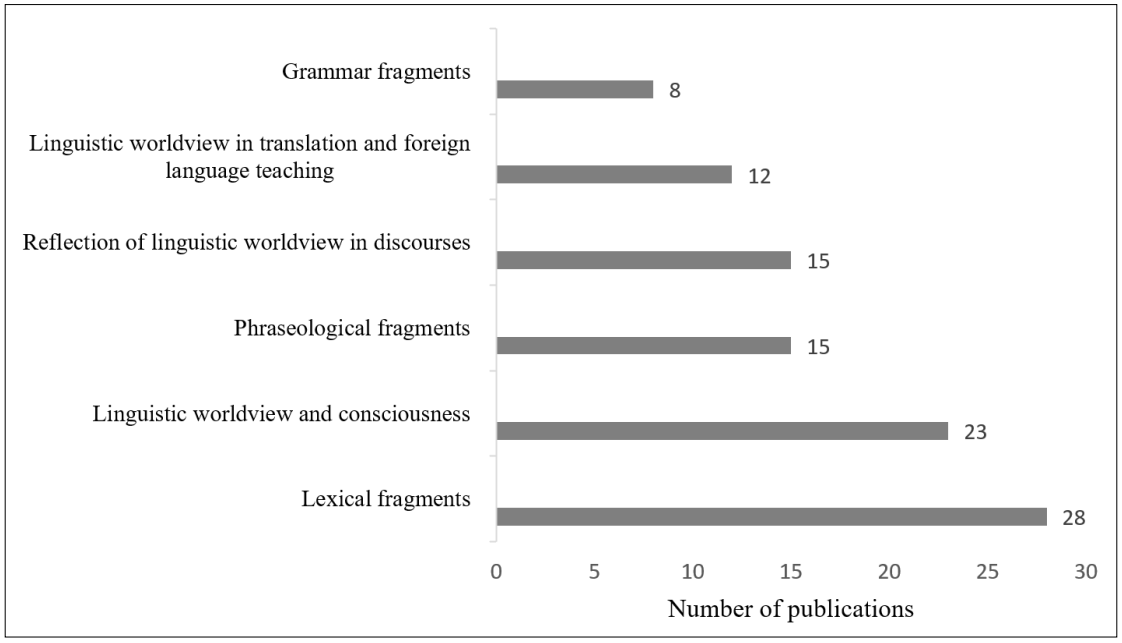
To form a foreign-language worldview in the minds of students, researchers offer various methods (4 papers). For example, they suggest that learning anthropomorphic vocabulary helps students understand English metaphorical transfers and form a secondary linguistic worldview (Pesina et al., 2019). Also, special attention is paid to the linguistic worldview in teaching a mother-tongue (Mikołajczuk, 2015; Pacovska, 2015; Wierzbicka-Piotrowska, 2015).

Thus, the present scoping review showed that linguistic worldview implies language-entrenched interpretation of reality, and it tends to comprise several fragments represented in lexico-semantic word groups, phraseological units, grammar forms, syntactic structures, and so forth. Besides, this scoping review revealed the areas of research, the comprehensively studied issues as well as demonstrated the statistics on the frequency of related publications, the years of the highest propagation of the research interest and the journals that published these papers. This review allows for concluding that linguistic worldview still remains a challenging research area.

Figure 4 demonstrates six research directions mentioned above with the number of publications within the period 2004-2022.

Figure 4

Areas of Research and the Number of Publications.



DISCUSSION

The given scoping review aims to determine, summarise and present a succinct overview of the existing knowledge on linguistic worldview in order to advance understanding of the linguistic worldview domain. For a detailed outline of the linguistic worldview research domain, special emphases were given to fundamental and empirical foundations of the literature on the issue published to date. By analysing the received scientific findings on linguistic worldview, the present review elucidates how the researchers' interest in national linguistic worldviews across Europe and Asia has changed for the last 18 years, what fragments of linguistic worldview deserved researchers' interest as well as methods applied by researchers for modelling the linguistic worldview. Apart from that, the authors traced the topic relevance rate over years, the year of 2004 appearing to be the outset of linguistic worldview investigation. This review has highlighted different aspects of the literature in question, endowing statistical information.

As a result of research, it was found that only two studies (Sartini & Ahimsa-Putra, 2017; Hayduk, 2018) have analysed the literature dealing with linguistic worldviews. The paper by Haydruk briefly reviews the main philosophical ideas by Osgood, Wierzbicka, Palmer and Budagov concerning linguistic worldview. Sartini and Ahimsa-Putra confined their review area to the studies dealing with worldviews in general, with special attention to the implemented research paradigms such as structuralism, evolutionism, functionalism, ethno-science and so forth. Since worldview is regarded as a multidimensional phenomenon, the authors of the aforementioned studies investigated the worldview categories which included categories of time and space, causality, the self and the other, the natural and supernatural, and the sacred and profane. However, the authors of the present study focus on the linguistic aspect of the issue and give special attention to the fragments of linguistic worldview and the most frequently applied methods of modelling the linguistic worldview.

The research is based on 116 sources on linguistic worldview which include empirical and theoretical studies published in academic journals and books within the period 2004-2022. The application of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) allowed the authors to carefully select the literature and ensure relevance of the sources. As a result, it was found that empirical studies significantly prevail over theoretical ones and the peak of researchers' interest fell onto the years 2018 – 2019. The scientists from Russia and Poland were the most active contributors to linguistic worldview research out of 9 countries, which were observed for the purposes of the present study.

Also, this review revealed that the linguistic worldview area has seen an increase of interest in modelling various linguistic worldview fragments, which seems to be accounted for by several factors: first, this topic is align with the current anthropocentric paradigm; second, this area of research goes far beyond linguistics proper and overlaps with those traditionally developed by such sciences as ethnopsychology and ethnolinguistics. By applying VOSviewer software and texts

inspection the authors identified the linguistic worldview fragments represented in various universal concepts such as FEMALE, FAMILY, BOUNDAY, MONEY, etc; the system of language, i.e. vocabulary and phraseology; discourses, i.e. poetic, literary, religious, scientific, mass media and advertising discourses, and Internet communication.

The research found that the most frequently used methods of modelling the linguistic worldview are inter-complementary and mutually reinforcing and often used in the same paper. These methods include: (1) *associative psycholinguistic experiment* used for the purpose of comprehensive study of linguistic consciousness, its structuring and modelling. The analysed studies found out the verbal associative connections of individuals formed on the basis of their previous experience; (2) *conceptual analysis* aiming to identify the culturally distinguished features of significant concepts; (3) *a comparative method* designed to identify common and national-specific features in the linguistic worldview of different linguistic communities.

CONCLUSION

A national linguistic worldview reflects a wide diversity of the given language speakers' life aspects. The linguistic representation of a person's experience and knowledge is regulated by the system of each specific language, including the methods of conceptualisation and categorisation adopted in it. It has demonstrated that this research area is being actively studied in many countries, with special attention paid to the following linguistic worldview fragments: the language system - lexical, phraseological, and grammar fragments, reflection of linguistic worldview in different discourses, linguistic worldview in translation and teaching. As for directions of further research, there are large avenues for analysis of unexamined or underresearched national linguistic worldviews and their fragments; also, comparative studies could involve more languages, with special focus on the languages belonging to different language families. Apart from that, given that all the reviewed works deal with the verbal representation of the linguistic worldview, it seems quite reasonable to investigate non-verbal means of communication in terms of national linguistic worldview reflection.

DECLARATION OF COMPETING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Tatyana Baranovskaya: supervision, writing - review and editing.

Marina Antonova: Conceptualisation, writing - original draft presentation.

Anna Zakharova: software, data curation, methodology.

Stanislav Li: visualisation, investigation.

REFERENCES

Abdullayeva, F., Li, S., & Proskurina, A. (2019). Teleut Labor and Life. In *Handbook of Research on Ecosystem-Based Theoretical Models of Learning and Communication* (pp. 294-307). IGI Global. https://doi.org/10.4018/978-1-5225-7853-6.ch017

Ajdarkhan, A. A. (2018). Correlation between language worldview and mentality in modern linguistics. *Emerging Scientist* (20), 436-439. https://moluch.ru/archive/206/50374/

Alefirenko, N., & Stepanova, K. (2015). Linguistic World Picture: Logoepisteme-domain-linguoculturеme. *Theory and Practice in Language Studies*, *5*(1), 1-6. https://moluch.ru/conf/phil/archive/235/12592/

Alefirenko, N., Dekhnich, O, Romashina, O., & Avdeeva, O. (2019). The ethno-cultural essence of linguistic consciousness. *Gêneno & Direito*, 8(7), 285-300.

Alyoshin, A. (2018). Swedish similes describing a person's movement as part of the linguistic worldview (in comparison with the Russian language material). *Scandinavian Philology*, 16(2), 193–207. https://moluch.ru/archive/206/50374/

Antipina, E. (2020). I.A. Bunin's Linguistic Worldview in an Extreme Situation (on manipulation potential of the diary «Cursed Days»). *Theoretical and Practical Issues of Journalism*, 9(2), 387-395. https://doi.org/10.2307/309957

Araeva, L., Artemova, T., Bulgakova, O., Kreydlin, G., & Obraztsova, M. (2015). Propositional frame-based description of fragments of Teleuts' linguistic worldview. *Review of European Studies*, 7(6), 295-301.

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International journal of social research methodology*, *8*(1), 19-32. https://doi.org/10.1080/1364557032000119616

Baimuratova, U., & Korobeynikova, A. (2020). Verbalization of the concept woman in the Russian and English orthodox discourse. In *Proceedings of the Philological Readings (PhR 2019)* (pp. 705-716). https://doi.org/10.1016/j.sbspro.2016.12.052

Barczewska, S., Wileczek, A., & Barańska, P. (2016). 'School? You go because you have to'. The linguistic worldview of 'school' in Polish and American teen Internet discourse. *L1 Educational Studies in Language and Literature*, *16* (Running Issue). https://doi.org/10.17239/L1ESLL-2016.16.01.08

Bartminski, J. (2012). Viewpoint, perspective, and linguistic worldview. *Aspects of Cognitive Ethnolinguistics* (pp. 76-87). Sheffield, Oakville: Eqinox. https://doi.org/10.1075/rcl.8.2.07ber

Bartmiński, J. (2013). The cognitive definition as a text of culture. *The Linguistic Worldview: Ethnolinguistics, Cognition, and Culture* (pp. 161-180). London: Versita.

Bartmiński, J. (2015). The semasiological and onomasiological perspectives in studying the linguistic worldview. [Perspektywa semazjologiczna i onomazjologiczna w badaniach językowego obrazu świata]. Poradnik Jezykowy, (1), 14-29.

Bartmiński, J. (2019). Borrowings and the Linguistic Worldview or How to Domesticate Foreignness. *Languages–Cultures–Worldviews* (pp. 237-258). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-28509-8_10

Bazarova, L., & Gilyazeva, E. (2018). Representation of The Concept "Woman" In English, Russian and Tatar Linguocultures (on The Material of Proverbs and Sayings). *Modern Journal of Language Teaching Methods*, 8(11), 185–190.

Bekasova, E. (2019). Peculiarities Of Linguistic Worldview Transformation on The Territory of Later Settling. In *The European Proceedings of Social and Behavioural Sciences (EpSBS)* (pp. 548–553). https://doi.org/10.15405/epsbs.2018.04.02.78

Berghout, A. (2006). Concept of Worldview Between Assumption and Truth (Observations on Selected Western and Muslim Views). *Jurnal Usuluddin*, (23), 117-138.

Bolshakova, N., & Gladkova, O. (2020). The Element of Fire in Dialectal Linguistic Worldview. In *Digital Future Economic Growth, Social Adaptation, and Technological Perspectives* (pp. 343-352). Springer, Cham.

Bulgakova, N., & Sedelnikova, O. (2015). Verbalization of the concept "Machine" in the Linguistic World Views of Russian and English Speakers. *Procedia - Social and Behavioral Sciences*, (215), 287–292. https://doi.org/10.1016/j.sbspro.2015.11.636

Carisio, A. (2020). The Linguistic Image of the World and Image Schemas: an attempt at their delimitation and comparison. *Complutense Journal of English Studies*, (28), 207-216. https://doi.org/10.5209/cjes.67041

Chakyroglu, A., & Suiyerkul, B. (2014). Representation of the concept "Hospitality" in the Kazakh language. *Procedia - Social and Behavioral Sciences*, (136), 124–128. https://doi.org/10.1016/j.sbspro.2014.05.301

Chirich, I.& Shtukareva E. (2019). Hospitality Vocabulary in the Russian Linguistic World Image. *Sovremennye Issledovaniâ Social'nyh Problem*, 11(5), 255–267.

Christmann, H. H. (1967). *Beiträge zur Geschichte der These vom Weltbild der Sprache*. Wiesbaden: Akademie der Wissenschaften und der Literatur.

Churilova, I. (2020). Representation Of Theatre Metaphors In The English Linguistic Worldview. In *European Proceedings of Social and Behavioural Sciences* (pp. 196-206).

Dalabayeva, A. (2014). Linguistic worldview: Of the world: Linguistic and national particularities. *Life Science Journal*, *11*(SPEC. ISSUE 5), 149-152.

Daudt, H. M., van Mossel, C., & Scott, S. J. (2013). Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. *BMC medical research methodology*, 13(1), 1-9.

Daulet, F., Zeinolla, S., Omarova, M., Smagulova, K., Orazakynkyzy, F., & Anuar, S. (2019). Somatic cultural code and its role in the Chinese linguistic worldview (based on the concepts of "face" and "heart"). *Humanities and Social Sciences Reviews*, 7(4), 703-710. https://doi.org/10.18510/hssr.2019.7490

Davaa, U. (2015). Experimental study of the linguistic worldview of Russians and Mongolians in comparative aspect (based on subjective defining of words denoting pets). *Bulletin of Kemerovo State University*, 4(4), 233–239.

Degel, L. (2010). The conceptual roles of intellectual abilities in the worldview of Estonian. *Eesti Rakenduslingvistika Uhingu Aastaraamat*, (6), 51-66.

Dziuba, E., & Podsukhina, V. (2019). On The Problem of Categorization of Artifacts in the Russian Linguistic Worldview. In *European Proceedings of Social and Behavioural Sciences*, (39), 27–33.

Erofeeva, V., Gilyazov, A., and Pilgun, M. (2019) Light as a Significant Concept of Medieval Religious Consciousness (As Exemplified in 'The Hagiography of Theodosius of the Kyiv Caves'). *Journal of Research in Applied Linguistics*, (10), 927–935. https://repository.kpfu.ru/eng/?p_id=233063&p_lang=2

Fefelova, G., Fatkullina, F., & Akhmatyanova, Z. (2018). The Concept of "Laughter" as a Form of Humorous Discourse (Using the Example of the Novel "12 Chairs" by I. Ilf and E. Petrov). *Modern Journal of Language Teaching Methods*, 8 (9), 307–316.

Gabbasova, A. & Fatkullina F. (2013). Linguistic worldview: main features, typology and functions. *Modern problems of science and education*, (4), 302-302.

Gasymova, G., Nelyubova, N., & Finskaya, T. (2022). The Role of the Diad PEACE-WAR Value in the Linguistic Worldview of the French Society Representatives. *RUDN Journal of Language Studies, Semiotics and Semantics*, 13(1), 156–173. https://doi.org/10.22363/2313-2299-2022-13-1-156-173

Gicala, A. (2013). The Linguistic Worldview and Conceptual Disintegration: Wisława Szymborska's Poem Identyfikacja and Its English Translation by Clare Cavanagh. *The Linguistic Worldview: Ethnolinguistics, Cognition, and Culture*, 61-75.

Gicala, A. (2021). Translating a worldview: Linguistic worldview in literary translation. *Linguistic Worldview(s): Approaches and Applications*. Routledge, 1-191.

Gieroń-Czepczor, E. (2020). National culture values in folk wisdom: The Indulgence (versus Restraint) Dimension in Polish proverbs. *SKASE Journal of Theoretical Linguistics*, 17(3), 59–77.

Głaz, A. (2022). *Linguistic Worldview (s): Approaches and Applications*. Routledge. https://doi.org/10.4324/9781003018803

Adam, G., Danaher, D. S., & Łozowski, P. (2013). The Linguistic Worldview. *Ethnolinguistics, Cognition, and Culture*. London: Versita.

Gordiienko-Mytrofanova, I., & Kobzieva, I. (2018). Concept "Holy Fool" in the Linguistic World-Image of the Russian-Speaking Population of Ukraine. *Psycholinguistics*, 24(1), 118–133. https://doi.org/10.31470/2309-1797-2018-24-1-118-133

Graneva, I. (2021). The pronoun ya ('i') in Russian phraseological units in the light of key ideas of the Russian linguistic worldview. *Vestnik of Northern (Arctic) Federal University. Series "Humanitarian & Social Sciences"*, (2), 42–51.

Grzegorczykowa, R. (2015). What is the linguistic worldview and why is it worth studying? *Poradnik Językowy*, (1), 7-13.

Gu, J. (2016). Humboldt's Views on Language, Translation and English Translation of Classics. Journal of Suzhou University of Science and Technology, 33(2), 97-102. https://doi.org/10.1515/9783110869484.3

Gulik, O. O. (2015). The Comparison of the Linguistic World-Image (the case of gender-marked cognitive metaphors with the names of animals as a component). *RUDN Journal of Language Studies, Semiotics and Semantics*, (1), 24-27.

Hayduk, N. (2018). The leading linguists research review concerning the national peculiarity of linguistic and conceptual world pictures reflection. *Scientific Journal of Polonia University*, *31*(6), 96-102.

Humboldt W. von. (1949) *Über die Verschiedenheit des menschlichen Sprachbaus und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechts*. Darmstadt: Claassen & Roether.

Humeniuk, I. (2021). Gender Markers in the English Picture of the World. *WISDOM*, 19 (3), 100–113. https://www.wisdomperiodical.com/index.php/wisdom/article/view/506

Jakubowicz, M. (2015). The linguistic worldview in etymological research. *Poradnik Jezykowy*, (1), 45-52.

Gao, J & Hua, Y. (2021). On the English Translation Strategy of Science Fiction from Humboldt's Linguistic Worldview - Taking the English Translation of Three-Body Problem as an Example. *Theory & Practice in Language Studies*, 11(2), 186–190. https://doi.org/10.17507/tpls.1102.11

Kalashnikova L. (2018). Linguistic worldview virtualization exemplified by K.S. Aksakov's story «Cloud». [*Aktual'nye Problemy Filologii i Pedagogicheskoj Lingvistiki*], (1), 51–61.

Khayrullina, R., Fatkullina, F., Morozkina, E., Saghitova, A., Suleimanova, A., & Vorobiov, V. (2016). Language personality in the bilingual context of the national linguistic world pictures. *Man in India*, 97(16), 17-26.

Khuzina, E., & Mukhtarova, R. (2018). The Lexico-Semantic Analysis of the Common and Various in World of Gender Picture in the English and Tatar Linguistic Cultures. *Journal of History Culture and Art Research*, 7(4), 214–222.

Kolmogorova, A. (2015). Prototypical linguistic worldview and word semantics Tacito consensus: Experimental studies. *Tomsk state university journal of philology*, 38(6), 29-42.

Komarova, A., Krylova, O., Gimranova, T., & Kolosova, E. (2017). Concept "Wedding" Existence in the Russian Linguoculture. *Ad Alta-Journal of Interdisciplinary Research*, 7(2), 9–12.

Konovalova, M. & Basova, V. (2021). Phraseological units as a way to compare Russian and Swedish linguistic worldviews. *Studia Humanitatis*, (18), 43-49.

Korovina, S., Pushkina, A., Krivoshlykova, L., & Ilina, A. (2020). Linguistic and cultural worldview of modern bilingual Chinese American writers. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 12(1). 1-16.

Kostina, N., Zerkina, N., & Pesina, S. (2014). Abbreviational Worldview as Part of Linguistic Worldview. Procedia - Social and Behavioral Sciences, (192), 703-709. https://doi.org/10.1016/j.sbspro.2015.06.065

Kotova, M., & Raina, O. (2020). Towards a linguistic vision of the world at the paremiological level of language. *Vestnik of Saint Petersburg University. Language and Literature*, 17(3), 487–504.

Kozlova, T. (2016). The Theory of Linguistic Worldview by Wilhelm Von Humboldt and Its Connection with the Culture of Sign Language. In *Proceedings of the 2016 International Conference on Arts, Design and Contemporary Education* (pp. 177-180).

Krylova, T. (2019). Distinguishing between animals and other living beings in the Russian linguistic worldview. [Voprosy Jazykoznanija], (2), 31-46.

Kurbatova, T. (2017). Basic features of conceptual sphere "geopolitics" in modern English-speaking worldview. *At the Crossroads: Challenges of Foreign Language Learning*, 37-51.

Kuryłowicz, B., & Szamryk, K. K. (2021). Linguistic and cultural perceptions of selected occupations in eighteenth-century proverbs (on the basis of Nowy dykcjonarz by Michał Abraham Troc). *Studia z Filologii Polskiej i Słowiańskiej*, 56.

Kushmar, L. (2019). Interrelation between conceptual and linguistic worldviews (based on the economic lexical item 'CLIENT'). *Psycholinguistics*, 25(2), 164-180. https://doi.org/10.31470/2309-1797-2019-25-2-164-180

Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. Chicago: U Of Chicago P.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago and London.

Lasitsa, L. (2020). Stereotypes of Russian Students About Great Britain and France. In *European Proceedings of Social and Behavioural Sciences*, (83), 261–270.

Lenart, I., & Markovina, I. (2021). Differences of kindergarten children's linguistic picture of the world: focus on Hungary, Russia, and Laos. *Heliyon*, 7(2). 1-17. https://doi.org/10.1016/j.heliyon.2021.e05940

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: advancing the methodology. *Implementation science*, *5*(1), 1-9. https://doi.org/10.1186/1748-5908-5-69

Levisen, C. (2019). The cultural semantics of untranslatables: Linguistic worldview and the Danish language of laughter. In *Languages–Cultures–Worldviews* (pp. 319-346). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-28509-8_13

Li, S., Araeva, L., & Guo, X. (2018). Reflection of national and cultural specifics in the text of advertising (on material of texts of advertising of the Chinese and Russian building companies). *MATEC Web of Conferences*, (212), 1–7.

Litvishko, O. M., Shiryaeva, T. A., Tikhonova, E. V., & Kosycheva, M. A. (2022). Professional discourse: The verbal and visual semiosis interplay. *Research Result. Theoretical and Applied Linguistics, 8*(1), 19-40. https://doi.org/10.18413/2313-8912-2022-8-1-0-2

Makarova, A., Gazilov, M., & Gozalova, M. (2018). The Issue of Unity and Diversity of Linguistic Worldview (On the Basis of Comparative Research of Idiomatic Phrases of French and Russian Languages). *International Multidisciplinary Scientific Conference on Social Sciences & Arts SGEM*, (5), 617–623.

Markelova, T.V., Petrushina, M.V., & Savelyeva, A.A. (2016). Axiological aspect of linguistic view of the world in the modern education environment. *SHS Web of Conferences*, (29), 02027.

Martynov, M. (2014). Specific features of anarchic denial in context of Russian linguistic worldview. Politeia, (88), 14-23.

Merzlikina, O. (2021). Zoomorphic metaphors "Livestock" in Russian and Galicia language pictures of the world. [*Vestnik Tomskogo Gosudarstvennogo Universiteta, Filologiya*], (71), 114-132.

Mierzwińska-Hajnos, A. (2010). The Linguistic Worldview Revisited. A Cognitive Analysis of Plant Terms. *Poznań Studies in Contemporary Linguistics*, 46(4), 457–479. https://doi.org/10.2478/v10010-010-0023-5

Mierzwińska-Hajnos, A. (2013). Jerzy Bartmiński's linguistic worldview meets the western cognitive tradition: The semantics of Polish and English plant names. *The linguistic worldview: Ethnolinguistics, cognition, and culture* (pp. 387-412). https://doi.org/10.2478/9788376560748

Mikołajczuk, A. (2015). The linguistic worldview in the polish language teaching versus teaching contents in the core curriculum and selected school handbooks. *Poradnik Jezykowy*, (1), 53-68.

Milovanova, M., & Matrusova, A. (2020). Thoughts on boundaries in the Russian linguistic consciousness: background and traditions. *Current issues of the Russian language teaching XIV*, 174–181.

Mirzoeva, L. (2019). Axiological component of the linguistic worldview (on the material of "Hagiography of Protopope Avvakum"). *Neophilology*, 29-34.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, *151*(4), 264-269.

Narushevich, A., & Bak, H. (2021). The category of animacy-inanimacy in the Russian language and the linguistic worldview. *E3S Web Conf*., 273, 11025.

Novoseletska, S., Shapran, N., & Musiichuk, T. (2020). Conceptual and linguistic worldview in modern linguistics. *Asia Life Sciences*, Supp22(2), 287-300.

Obraztsova, M., Denisova, E. & Kereksibesova, U. (2019). The Symbolism of Archaic Rites, Signs and Superstitions of Teleuts. *Handbook of Research on Ecosystem-Based Theoretical Models of Learning and Communication*, E.A. Railean (Ed.). Hershey, PA: IGI Global, 277—293.

Olenev, S. V., Araeva, L. A., & Bulgakova, O. A. (2019). Teleuts' Family and Kinship Ties: Socio-Demographic Background and Linguistic Analysis. *Handbook of Research on Ecosystem-Based Theoretical Models of Learning and Communication* (pp. 308—323). Hershey, PA: IGI Global. https://doi.org/10.22363/2313-2299-2019-10-4-841-859

Olomskaya, N., Patyukova, R., Kudinova, T., Demina, L., & Cherkasova, M. (2018). Lexico-semantic analysis of the emotion "grief and suffering": Crosscultural overlapping of linguistic worldviews (on the material of Russian and English languages). *XLinguae*, 11(1), 242-250.

Pacovska, J. (2015). Cognitive-Oriented Teaching of the Mother Tongue. *2nd International Multidisciplinary Scientific Conference on Social Sciences and Arts (SGEM 2015)*, (2), 159–166. http://sgemsocial.org/

Panic, N., Leoncini, E., De Belvis, G., Ricciardi, W., & Boccia, S. (2013). Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PloS one*, *8*(12), e83138. https://doi.org/10.1371/journal.pone.0083138

Pajdzińska, A. (2013). The Linguistic worldview and literature. *The Linguistic Worldview: Ethnolinguistics, Cognition, and Culture* (pp. 41-59). https://doi.org/10.2478/9788376560748

Pesina, S., Kiva-Khamzina, Y., & Rubanova, N. (2019). Formation of an idea of the secondary linguistic and conceptual worldview of bachelor's students studying a foreign language. *Perspektivy Nauki i Obrazovania*, 37(1), 180-189.

Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation*, *13*(3), 141-146.

Pivovar, Y. (2019). The key concept as a structuring base of Russian and Belarusian linguistic worldview. *Russian Language Studies*, 17(2), 243-254.

Pospelova, N., Achaeva, M., & Koroleva, E. (2021). Lingvocultural Types in the British Language Picture of the World. *Propósitos y Representaciones*, 9(SPE2), e1021. https://doi.org/10.20511/pyr2021.v9nSPE2.1021

Qin, J., & Zeng, W. Humboldt's View on Translation and its Enlightenment to the English Translation of Classics (2010) Journal of University of Electronic Science and Technology of China, 12(4), 77-81.

Revenko, I. & Osetrova, E. (2019). Two "Characters" in the Russian Linguistic Worldview: "We" and "They". Journal of Siberian Federal University. *Humanities & Social Sciences*, 1-16. https://elib.sfu-kras.ru/handle/2311/128352

Rudenka, A., & Fang, X. (2018). Nomination of mental illnesses in languages of different types. *In SKASE Journal of Theoretical Linguistics* [online], 15(3), 155–159.

Sadykova, A., & Kayumova, D. (2014). The correlation between linguistic and conceptual worldviews: The role of metaphor. *Life Science Journal*, 11(6), 552-555.

Salnikova, V. & Asanbaeva, S. (2017). Child's linguistic worldview: experimental studies. *Filologické vědomosti*, (2). 23-27.

Samedov, J., & Gasanova, M. (2017). Concepts work and laziness in the archi language worldview. *Voprosy Kognitivnoy Lingvistiki*, (1), 39-44.

Sapir, E. ([1929] 2000). Status of linguistics as a science. In Edward Sapir*, Culture, Language and Personality* (pp. 65-77). Mandelbaum Berkeley and Los Angeles: University of California Press.

Sartini, S., & Ahimsa-Putra, H.S. (2017). Preliminary study on worldviews. *Humaniora*, 29(3), 265-277. https://doi.org/10.22146/jh.29690

Sawaniewska-Mochowa, Z. (2018). The names of values in the linguistic worldview of Lithuanians. *Acta Baltico-Slavica*, (42), 308-315. https://doi.org/10.11649/abs.2018.017

Shapovalova, E. (2019). The Concept "Money" In the French Linguistic World Image. In *European Proceedings of Social and Behavioural Sciences*, (39), 426–431. https://doi.org/10.15405/epsbs.2018.04.02.60

Shatilova, L. M., Borisova, V. V., & Kasatkina, O. A. (2018). Representation of the linguistic and cultural concept "lie" in the French and Russian language picture of the world. *Opcion*, 34(85-2), 257-276.

Shkuropatskaya, M., & Undarmaa, D. (2016). National linguistic world image as a component of linguistic consciousness of a Russian and Mongolian language personality (a comparative aspect). *Vestnik Tomskogo Gosudarstvennogo Universiteta, Filologiya*, (1), 80-95.

Shkuropatskaya, M., & Davaa, U. (2018). National linguistic world-image in mental-speech activity of Mongolian and Russian native speakers (comparative study). *Sibirskii Filologicheskii Zhurnal*, (2), 141-156.

Shulgina, E., & Fang, Y. (2014). The Concept "Family" in the Russian and Chinese Linguistic Views of the World. *Procedia - Social and Behavioral Sciences*, (154), 162–169. https://doi.org/10.1016/j.sbspro.2014.10.129

Shumilo S. (2021). The Category of Tense in a Medieval Human's Outlook: Influence of Linguistic Worldview on Verb Morphology. *Zenodo*, 134-142. https://doi.org/10.1016/j.sbspro.2015.06.065

Sipko J. (2018) The Fragments of Linguistic worldview of Immigration in the Countries of the Visegrad Group. *Political Linguistics*, (4), 167-174.

Sosnowski, W., & Tymoshuk, R. (2017). On The dictionary of active Polish and Ukrainian phraseology. Contrastive linguistics and culture. *Cognitive Studies | Études cognitives*, Article 1317. https://doi.org/10.11649/cs.1317

Stria, I. (2015). Linguistic worldview of Esperanto: A questionnaire method. *Empirical methods in language studies*, 41-51.

Suleimenova, E. (2013) Reminiscences about the world view/linguistic worldview. Bulletin of Kaznu. Philologicalseries, 3(143), 121-125.

Szarlej, J. (2020). The linguistic picture of the woman and the man in the prophetic books of the Old Testament. *Świat i Słowo*, 34(1), 359–376.

Tagirova, T. (2018). A culture-specific "linguistic picture of the world" in the organization of scientific discourse in a foreign language. In *The International Science and Technology Conference" FarEastCon"* (pp. 19-26). Springer, Cham. https://doi.org/10.1007/978-3-030-18553-4_3

Tokarski, R. (2016). From the linguistic picture of the world to the pictures of the world in language. *Jezyk Polski*, 96(2), 28-37. https://doi.org/10.15688/jvolsu2.2019.1.1

Tóth, S. J. (2018). The image of the world in Slovak and Hungarian grammaticalised categories. *Yearbook of Finno-Ugric Studies*, 12 (4), 15–27.

Urszula, K. (2016). The linguistic worldview of the dog in tombstone inscriptions in virtual pet cemeteries. *Poradnik Jezykowy*, (9), 48-57.

Ustinova, O. (2014). Canadian Literature as a Linguistic Reflection of the Worldview of French and English Canadians. *RUDN Journal of Language Studies, Semiotics and Semantics*, (1), 74–82.

Valkova, Y. (2018). The linguistic worldview of contemporary Israeli writers (E. Kishon, A.B. Yehoshua, A. Oz, M. Shalev, D. Grossman) in Russian translation: Estrangement and translation strategies. *Journal of Siberian Federal University - Humanities and Social Sciences*, 11(1), 155-170.

Weisgerber, L. (1930). Die Zusammenhänge zwischen Muttersprache, Denken und Handeln. In *Zeitschrift fur deutsche Bildung,* (6), 57-72.

Weisgerber, L. (1931). Sprache. In *Handbuch der Soziologie*. In Verbindung mit G. Brieis, F. Eulenburg, F. Oppenheimer, W. Sombart, F. Tonnies, A. Weber, L.V.Wiese. Hrsg. von A. Vierkandt (pp. 592-608). Stuttgart, Enke.

Wierzbicka-Piotrowska, E. (2015). Methodology of the linguistic worldview in school teaching. *Poradnik Jezykowy*, (1), 69-78.

Wyrwa A. (2017). "Coming out" or "wychodzenie z szafy": A comparative analysis of linguistic worldviews in British and Polish press. *Półrocznik Językoznawczy Tertium*, 2(2), 98–118.

Xiu, W., Jiang, T. Translation of Science Fiction from the Perspective of Language Culture-Communication--A Case study of Liu Kun's Translation of Folding Beijing (2018) *Chinese Science and Technology Translators Journal*, 31 (1), 46-49.

Yergaliyev, K., Asanbayeva, Y., Amrenov, A., & Shaharman, G. (2015). Linguistic view of the world and peculiar features of its reflection in newspaper headlines. *Mediterranean Journal of Social Sciences*, 6(5), 264-268. https://www.richtmann.org/journal/index.php/mjss/article/view/7487

Yuldashev, A. (2020). Idiomaticity of the Uzbek linguistic worldview. *Voprosy Kognitivnoy Lingvistiki*, (1), 130-135.

Zakharov, M., Prokopieva, S., & Zakharova, A. (2021). Metaphors and linguistic image in the analysis of phraseological units of the archaic Olonkho narrative. *Revista de la Universidad del Zulia*, 12 (34), 572–584. http://dx.doi.org/10.46925/rdluz.34.31

Zawisławska, M. (2015). The function of metaphor in the reconstruction of the linguistic worldview on the example of the use of metaphors in the language of vintners. *Poradnik Jezykowy*, (1), 79-88.

Zinken, J. (2004). Metaphors, stereotypes, and the linguistic picture of the world: Impulses from the Ethnolinguistic School of Lublin. *Metaphorik. de*, (7), 115-136.

https://doi.org/10.17323/jle.2022.13785

The Routledge Handbook of World Englishes: A Book review

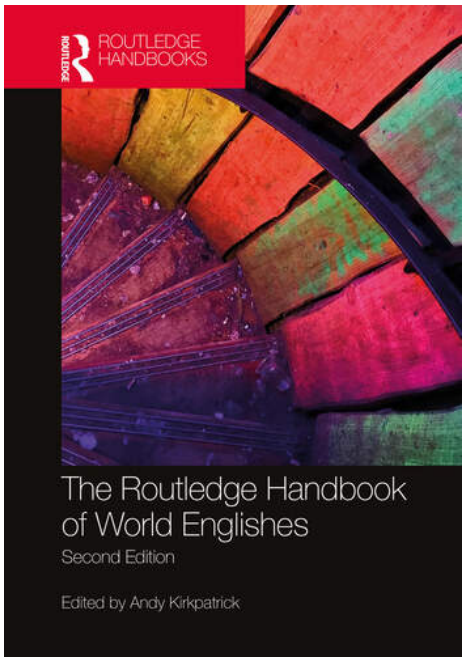
Fadhila Yonata [1], **Dwi Rukmini** [2], **Suwandi** [3], **Sri Wuli Fitriati** [2]

[1] STAIN Sultan Abdurrahman Kepulauan Riau, Semarang, Indonesia

[2] Universitas Negeri Semarang, Semarang, Indonesia

[3] Univeristas PGRI Semarang, Semarang, Indonesia

Kirkpatrick, A. (2021). The Routledge Handbook of World Englishes (2nd ed.). Routledge. ISBN 9780367144395



Through the years, English has increasingly been spoken by millions of speakers in the world as a second and foreign language to communicate across countries. As it was natively used by, following categorization conceptualized by Braj Bihari Kachru (1990), inner circle countries (English as L1), the speakers of outer circle (English as L2) and expanding circle (English as a foreign language, EFL) have outnumbered the native speakers (Kirkpatrick & Sussex, 2018). The cultural, geographical, and physical distinctions of speakers worldwide may shape and transform linguistic features (Kirkpatrick & Lixun, 2021). It leads to varieties of English. Bearing this in mind, through the lens of the use of English worldwide shaped by his upbringings, Andy Kirkpatrick, experienced living in the U.K., Malaysia, and Australia edited a sheer handbook of world Englishes by compiling 40 cutting-edge chapters capturing the varieties of English around the world, from its home origins involving U.K., U.S., Canada, Australia, and New Zealand to outer and expanding circles (European, African, and Asian countries) and also comprises varieties across social disciplines (culture, education, and business). As the successor of its preceding edition, the second edition attempted to address extensively the up-to-date issue on the role of English including socio-cultural, language literature, and pedagogical concerns and welcome readers across disciplines.

Over the last decades, the notion of World Englishes (WE) has sparked generative discussions on how to anticipate and compensate for the diversity of linguistic features of speakers. This volume, comprising six sections, once again impressively highlights the role of English in people's life including historical perspectives, regional varieties of English as 'new English' from across the world, recent and emerging trends of English from a linguistics perspective as well as the pedagogical implications for language teaching and the prediction of the future of Englishes. All the chapters provide reflections and call for language practitioners and educators to revisit the standard of English based on the context in which the language is used. This laudable book is good food for thought for those, not limited to, who is in charge of policymakers, students of applied linguistics, and those involved in the education domain (curriculum designers, classroom planners, and teachers).

The eight chapters in the first section are grouped under the heading "Historical perspectives and traditional Englishes". The editor positions the pertinent cave-

Citation: Yonata F., Rukmini D., Suwandi, & Fitriati S. W. (2022). The Routledge Handbook of World Englishes: A Book review. *Journal of Language and Education, 8*(3), 167-169. https://jle.hse.ru/article/view/13785

Correspondence:
Fadhila Yonata,
fadhilayonata92@gmail.com

Received: January 21, 2022
Accepted: September 17, 2022
Published: September 30, 2022

at in order, beginning with the opening chapter by Daniel R. Davis, entitled "Standardized English: The history of the earlier circles", which documents the periodization of English transformation from old-time to modern English era. Rooted in Kachru's classification of World Englishes (Kachru, 1990), English in the inner circle countries is the main theme of the following chapters. To begin with, in the scope of British English, the reports on language development are focused on grammatical variation (David Britain, Chapter 2) and phonological innovation (Gerard Docherty, Chapter 3) in contemporary English in England. Elucidating how English is spoken in some cities of England, Britain reveals the profusion of non-standard English in daily interaction while Docherty figures out the current changes of pronunciation in young speakers. Further, Raymond Hickey, the author of chapter 4, stipulates the current trends of English in Ireland. Bill Kretzschmar discusses American English while Stephen Levey focuses on the use of English in Canada. In the next chapter, Kate Burridge discovers "English in Australia" and Margaret Maclagan and Paul Warren report English in New Zealand, looking at how those varieties have their own characteristics each. One point should be noted that the topics covered in this section lead the readers to see the diversity of English in its home or inner circle countries as an accepted standard in communities, and likewise varieties of other countries are agreed as well to commence the World English paradigm.

The growing number of worldwide users have taken part in shaping and forming the expanding varieties of English. Section II bears the tell-tale heading 'Regional Varieties' and stipulates the array of chapters portraying English in India (Joybroto Mukherjee and Tobias Bernaisch), Sri Lanka (Dushyanthi Mendis and Harshana Rambukwella), Bangladesh (Obaidul Hamid and Iffat Jahan), East and West Africa (Hans-George Wolf), Maldives and Uganda (Christiane Meierkord), South Africa (Susan Coetzee-Van Rooy), China (Zhichang Xu), Philippine (Isabel Pefianco Martin), Singapore and Malaysia (Low Ee Ling), Japan (Yuko Takeshita), Slavic countries (Zoya Proshina), Caribbean (Hazel Simmons-McDonald), and Colombia (Adriana González). From the texts, one of the epistemological reasons for the transformation of English in making new varieties could be indicated by the impact of British imperialism and postcolonialism (except for Japan and Slavic countries). English has been institutionalized in the formal sphere, i.e., government, school, and legal documents. Thus, the uniqueness is possibly crafted due to regular usage.

Section III touches upon some of the most discussed themes and trends. The initial and informative chapter by Barbara Seidlhofer entitled "English as a lingua franca in the European context" discovers the uniqueness of English used in European countries which are classified as the expanding circle (Kachru, 1990). Edgar W. Schneider, in the next chapter, challenges the previous framework to categorize World English by promoting a new model of dichotomy namely postcolonial and non-postcolonial. As the imminent research trend on World Englishes, we personally believe this potential framework suitably replaces Kachru's seminal framework and accommodates the categorization of countries more clearly as the dynamic status of English in all countries. The following chapter by James McLellan explicates mixed codes of international practices. Farzad Sharifian points out the intertwined between language and culture. She emphasizes that the speakers of English from the outer and expanding circle "express and embody cultural conceptualizations of their speakers" (Kirkpatrick, 2021, p.437). With the domination of personal appraisals, Ha Jin, in chapter 26, reports individuals' perception of being "painful" speakers of English as a foreign language, since "the confinement of the colonial heritage" (Kirkpatrick, 2021, p.458). Ultimately, the section is closed with the chapter authored by Ahmar Mahboob headlined «World Englishes, social disharmonization, and environmental destruction". Grounded in the Sapir-Whorf Hypothesis (SWH) on language and cognition, language may influence its speakers' cognition and the existence of a language is determined by its speakers that the strength of the speaker impacts the strength of the language's status in society. For now, as non-native people massively adopt English, the domination may possibly hijack "the sociosemiotic inheritance of local communities" (Kirkpatrick, 2021, p.474) and it is somewhat a critical awareness of the role of English worldwide for global users.

The editor introduces a fruitful and potential issue for the foreseeable future, "Online English" by Mark Warschauer, Sharin Jacob, and Undraa Maamuujav at the outset of Section 4, «Contemporary contexts, functions and variables». The chapter by Catherine Nickerson examines the expansion of Englishes in the scope of business. In addition, the following chapters look at the use of English as a medium of instruction in higher education (Ernesto Macaro), popular culture (Andrew Moody), and call centers (Kingsley Bolton). The topical issues are also included namely "Translanguaging and multilingual creativity with English in the Sinophone world" by Tong King Lee and Li Wei which reveal that the massive impact of social media advances facilitates translanguaging as a medium of creating a new language. This dovetail with the means of new-standard innovation later in Chapter 37. Finally, in Chapter 34, Mario Saraceni, Britta Schneider, and Christine Bélanger focus on capturing the unchangeable role of English in Europe even after Britain's withdrawal from the European Union, known as "Brexit". They point out that "the roots of the English language reach deeply in the European sociocultural soil and are immune to Britain's status within or outside the European Union" (Kirkpatrick, 2021, p.589).

Section 5, one of the most pertinent parts of the volume, is a collection of five chapters that address "Debates and pedagogical implications". Chapters in this section provide the encyclopaedic coverage of language education starting with a variation of phonology by David Deterding. He emphasiz-

es the salient goal of language teaching for communication is intelligibility. Matin Dewey, the author of "English language teachers in context: Who teaches what, where and why?", brings out the thought-provoking issue of comparison between Non-Native Speaker Teachers (NNST) and Native Speaker Teachers (NST) in second and foreign language education. He calls for re-evaluation in considering the appropriate educators so that the context of targeted learners should be taken into account. David C.S. Li and Deyuan He inform the process of how an unconventional form becomes an innovation as a new variety. Anchored in the Bamgbose (1998) model, the key factors in reconstructing new non-standard English are demographic, geographical, authoritative, codification, and acceptability. One additional factor postulated in this chapter, in line with the growth of internet users globally, is "the popular choice of acrolectal English L2 users in cyberspace" (Kirkpatrick, 2021, p.637). Brian Tomlinson focuses on the English language used for proficiency tests. Despite the varieties of English, Standard English remains to be the norm as it deeply roots in the daily encounters of global speakers. Further goes a chapter written by Anna Mauranen, Carmen Pérez-Llantada, and John M. Swales headlined "Standard English is asserted as academic Englishes". Overall, the very chapter's titles hone in on language pedagogy and call for language researchers and educators to expand and delve into language education practices in every context.

Ultimately, the final section consists of only one single chapter. The volume ends up with controversy over the role of Englishes in the future authored by Alastair Pennycook. He poses the possibilities of English coming decades, "one English, many Englishes or no English» (Kirkpatrick, 2021, p.679). This becomes an intriguing question and puts this issue in a continuum as the result is highly related to global socio-political.

As readers in the field of foreign language education, we aspire to claim this volume to be a timely reference and have already considered it a breakthrough in the paradigm of English language teaching. The "native-speakerism", an old-fashioned ideology that uses inner-circle countries as a norm (Holliday, 2005), has been challenged for its longevity in the English education sphere. This edited volume has apparently set up a comprehensive portrait of WE across fields and provided special space to point out the pedagogical implication. The editor seems to anticipate the doubt and is against the drawbacks of the WE perspective. However, the book still has ample room for improvement. Albeit the comprehensive discussions, the theoretically informed discussion and adaptable framework of WE are limitedly discussed in the whole volume. The editor also states early on that "it is impossible for a Handbook such as this to encompass all varieties of World Englishes and all the developments and debates that surround the use of English in today's world" (Kirkpatrick, 2021, p.9). The editor along with the authors of each chapter has pursued a preferred line of thinking by taking the "glocal'" (both global and local) perspective into this volume which indicates the expectation of the WE framework applied as the fundamental consideration in English language education. Our final remark is that the importance of a volume such as this is of immense insight and a "must-read" for students as well as scholars in English language teaching and applied linguistics.

ACKNOWLEDGMENT

The first author would like to express the deepest gratitude to *Lembaga Pengelola Dana Pendidikan* (LPDP) for sponsoring his doctoral study, as well as supporting the publication of this article.

DECLARATION OF COMPETING INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Fadhila Yonata: Conceptualization, writing – original draft, writing – review & editing.

Dwi Rukmini: Supervision, validation.

Suwandi: Supervision, writing – original draft.

Sri Wuli Fitriati: Conceptualization, supervision, writing – review & editing.

REFERENCES

Holliday, A. (2005). The struggle to teach English as an international language. Oxford University Press.

Kachru, B. B. (1990). World Englishes and applied linguistics. World Englishes, 9(I), 3–20.

Kirkpatrick, A. (Ed.). (2021). The Routledge Handbook of World Englishes (2nd ed.). Routledge. https://doi.org/10.4324/9781003128755

Kirkpatrick, A. & Sussex, R. (2018). Introduction. In A. Kirkpatrick & R. Sussex (Eds.), English as an International Language in Asia: Implications for Language Education (pp. 1–12). Springer. https://doi.org/10.1007/978-94-007-4578-0

Kirkpatrick, A. & Lixun, W. (2021). Is English an Asian Language? Cambridge University Press. https://doi.org/10.1017/9781316471166

CONTENT

EDITORIAL

RESEARCH ARTICLES

PILOT STUDY

REVIEW PAPER

REVIEWS