Language Teaching Research Quarterly

2021, Vol. 22, 91-101

Language Bite-by-Bite: The Design of Second Language Learning Application Using Lexical Approach

Nuttanart M. Facundes*, Booncharoen Sirinaovakul

King Mongkut's University of Technology Thonburi, Thailand

Received *09 May 2021* ***Accepted*** *21 September 2021*

Abstract

The ability to use languages to communicate distinguishes humans from other living things. In language production, two sources of information play a crucial role: a conceptual (semantic) message and a sequence of syntactic states mapped onto an ordered sequence of words. This concept of language production applies to both the native language of the speaker and the second language. Therefore, one question arises as to why second language learners cannot equally master both their second language (L2) and first language (L1). In our research, we investigate second language learning within the linguistic and language production framework. We explain language competence by knowledge of paradigmatic and syntagmatic relations, and we can apply the lexical approach for language learning to reach such competence. With this basis, we have designed an educational web application for second language learning, *Bite-sized English*, that incorporates the learning of words and phrases hands-in-hands with grammar. Bite-sized English consists of two parts: "Take a Bite", where users learn lexical chunks from authentic texts, and "Lingua Kitchen", where learners can build their own lexicon and practice usage in contexts.

Keywords: *Lexical Approach, Second Language Learning, Paradigmatic, Syntagmatic, Language Production*

Introduction

The ability to use languages to communicate distinguishes humans from other living things. It is a fascinating human capability that allows us to express our thoughts, ideas, emotions in words. Humans acquire the ability to use their first language naturally. When investigating language production, Gordon and Dell (2003) propose a model of sentence production that maps two

nuttanart@cpe.kmutt.ac.th doi:10.32038/ltrq.2021.22.07

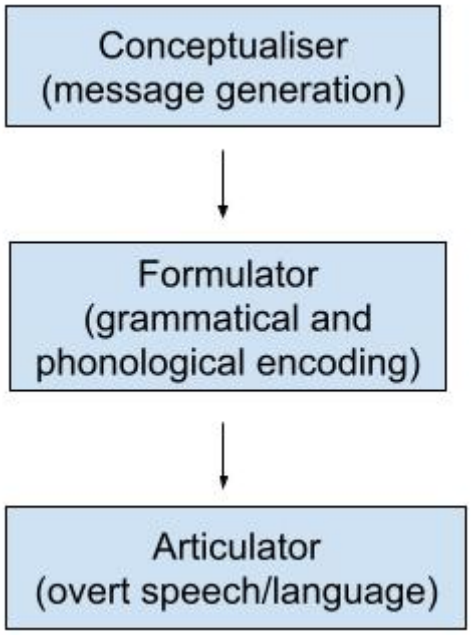
sources of information: a conceptual (semantic) message and a sequence of syntactic states onto an ordered sequence of words. This concept of language production applies to both the native language of the speaker and the second language. Therefore, one question arises as to why second language learners cannot equally master both their second language (L2) and their first language (L1).

Language Production and Language Competence: A Review

Language production is how humans produce a language, starting from having concepts to express to planning and realising in speech or written forms. Language production stages are mainly agreed to consist of 4 main components: conceptualisation, formulation, articulation, and self-monitoring. Here we will adopt Levelt's (1989) model as our research framework for language production. The model mainly consists of *a conceptualiser* consisting of message generation, *a formulator* consisting of grammatical and phonological encoding, and *an articulator* that produces overt speech/language. In this research, we focus on the conceptualiser and formulator components of the model.

Figure 1

Simplified Levelt's Components of Speech Production Model (1989)



Relevant to language production is the term *language competence* or *language proficiency*, which is internalised knowledge that native speakers may have of their language. This knowledge is not merely knowing how to speak a language, but it constitutes a speaker's linguistic competence, or what is acquired in learning a language, implicitly or explicitly

(Paradis, 1998), and used in language production and understanding. Language competence influences language production as Dell et al. (2021) claim that the language production model is adaptive and reflective, meaning that it can adapt to recent inputs and reflect the learning experience accumulated over time.

Paradigmatic and Syntagmatic Relations

According to the pioneer of modern Linguistics, Ferdinand de Saussure, language is a system whose members, or linguistic signs, are defined by their relations to the other members of the system (Saussure, 1959). Language thus comprises two-dimensional relations: *paradigmatic* and *syntagmatic*. Paradigmatic relations refer to elements that can replace one another or *choices*, while syntagmatic relations define combinatory possibilities or how elements combine in a sentence or *chains* (Culler, 1976). While syntagmatic relations concern sentence structures, word order, and relations among elements in a sentence, paradigmatic relations encompass semantic or word meanings and category information. To explain language production planning in terms of paradigmatic and syntagmatic relations, speakers have thoughts or concepts to express (semantics and categories) and arrange them into an utterance or a sentence.

Furthermore, according to Levelt et al. (1999), in speech production at the level of lemma selection in *formulator*, the words are presented holistically and not simply in forms and are activated incrementally. Consider the sentence "*The cat might chase the string.*" In speech production planning, when we focus on the target noun *"cat"* in the above sentence, the competitive nouns that might be activated in the formulator could be *"dog"* or *"cap"*, which are semantically or phonologically related to *"cat"*. This could be seen as paradigmatic competitions in speech production planning. Moreover, syntagmatic competitors such as the noun *"string"* could be activated. Evidence for paradigmatic and syntagmatic relations in language production planning can be seen from speech errors or slips of the tongue. A slip of the tongue is a one-time error in speech production planning, which results in the speaker producing an utterance different from the one originally intended. For the sentence mentioned above, the paradigmatic slip *"The dog might chase the string."* or the syntagmatic slip *"The string might chase the cat."* could happen. Slips of the tongue are a rich source for studying processes involved in speech production, particularly lexical selection and insertion, as well as units in which linguistic information is mentally stored. (Fromkin, 1973, 1980; Cutler, 1982; Stemberger et al., 1985; Dell, 1986; Jaeger, 1992). To summarise, we can see how lexical selection, paradigmatic and syntagmatic relation play a role in language production.

First and Second Language Learning

The previous sections describe how language production proceeds and its relation to language competence as well as syntagmatic and paradigmatic knowledge, which is acquired implicitly or explicitly and related to the terms *acquisition* and *learning*. In this paper, we focus on using the term *learning* with second-language learners. However, it is worth considering the difference between language acquisition and language learning (Limacher, 2017). The acquisition means

children acquire the ability to perceive and produce a language through a natural, subconscious process of verbal/nonverbal communication and are unaware of grammatical rules. This happens especially when children acquire their first language (L1) and is different from second language (L2) acquisition. In other words, first language acquisition is children's acquiring their native language, while second language acquisition is consciously learning a language after acquiring the first language. Acquiring first language does not require explicit instructions or education, but acquiring or learning second language requires instructions and/or intention to learn.

In summation, language acquisition means the process internally and naturally induced from the learners who are usually children and acquire their first language, while language learning implies learning (second) language with external and explicit instructions. However, there exists the term *second language acquisition* or SLA, which was mentioned earlier, and it means that the learners learn the second language possibly in a way that imitates the acquisition of their mother tongue. After we have explored the two terms: *acquisition* and *learning*, we prefer the use of *second language learning* as we will be discussing external tools for promoting second language learning without rejecting the term *second language acquisition* if it is used in the reference sources.

Nonetheless, in terms of the mechanism of language production discussed earlier, we assume it to be the same for both first language and second language. This supports by evidence that stages for L1 and L2 learners are the same, starting from labelling objects (one-word stage) and then having something to express about those objects (relations between words). What is in addition for L2 learning is "bridging" or tying L2 concepts to L1 concepts. What might also be different is that older children and adults apply the rules they consciously learn to self-correct their language production. Meanwhile, some of the rules are not fully internalised, resulting in lagging behind in L2 production.

We have seen the similarity and differences between L1 and L2 and the methods or processes to reach the capacity of using them. It should be noted that, according to Stephen Krashen (1993), learners should acquire second languages in the same way children learn their first. He proposed a natural approach of language teaching whereby language output is not forced but allowed to emerge after students have attended to large amounts of comprehensible language input. Learners should understand input language that contains structures superior to their current level of competence. Krashen defined a learner's current stage of studying language as "i" and the ideal level of input as i +1. The natural approach has influenced current theories of second language learning, such as the lexical approach by Lewis (1993; 1997), which will be discussed in the next section.

The Lexical Approach to Language Learning

In learning languages, learners are not learning words individually but in contexts. Suppose second language learners of English learn a new word: *"crew"*, they might find the translation of this word in L1 and think or learn of other words associated with it, such as *"ship"*, *"sailor"*, *"flight"* and *"airline"*. The learning of the word *"crew"* is not complete until it is placed in an

appropriate sentence or a context. Hence, the way the learners' lexicon is built up cannot be separated from the grammar and vice versa. That is, both paradigmatic and syntagmatic dimensions of language are relevant. Furthermore, according to Lewis (1993; 1997), the ability to produce lexical phrases or chunks such as *hot and spicy* or *make a point* is important in the lexical approach to language learning. A lexical chunk, which is also called "multi-word unit" or MWU (Wray, 2002), is an idiomatic or conventionalised phrase and is learned or processed as a whole. To illustrate this, the examples of lexical chunks and their categories from Thonbury (2019) are shown below:

• collocations (Examples: *wrong way, give way, the way forward*)
• fixed expressions (Examples: *by the way, in the way*)
• formulaic utterances (Examples: *I'm on my way; no way!*)
• sentence starters (Examples: *I like the way…*)
• verb patterns (Examples: *to make/fight/elbow one's way…*)
• idioms and catchphrases (Examples: *the third way; way to go*!)

The reason why lexical chunks should be units for language learning is because of their pervasiveness, especially in spoken language, as 45% of the words in a corpus of conversational English occurred in formulaic or bundle forms (Biber, 1999). In accordance with Lewis' recommendations (1993; 1997), lexical chunks can be applied to language teaching as follows:

1) Students' attention should be drawn to lexical chunks without teachers analysing and explaining them at the very beginning.
2) The importance of noticing and the practice of repeating tasks should be emphasised.
3) The traditional Present-Practice-Produce paradigm should be replaced by the Observe-Hypothesis-Experiment cycle.
4)Teaching activities should promote awareness-raising of lexical chunks in learners.

Furthermore, the organising of teaching and learning of lexical chunks are based on the following principles:

1) **Topic**: lexical chunks can be organised into topics.
2) **Situation**: physical situation or environment of the classroom.
3) **Collocation:** identifying chunks as collocations
4) **Notion**: comparing and contrasting lexical chunks
5) **Phonological chunking**: making use of intonational properties or formulaic speech of lexical chunks to make them easier to remember.
6) **Grammar**: although grammar is not the main focus, it is not discarded in learning lexical chunks.

Language Learning in Chunks: The Design of a Web Application

Nowadays, vocabulary learning is possible via many available software programs or applications. However, they focus on learning new words from prefabricated sets of words, check the learners' knowledge of words, or are for dedicated purposes of preparation for

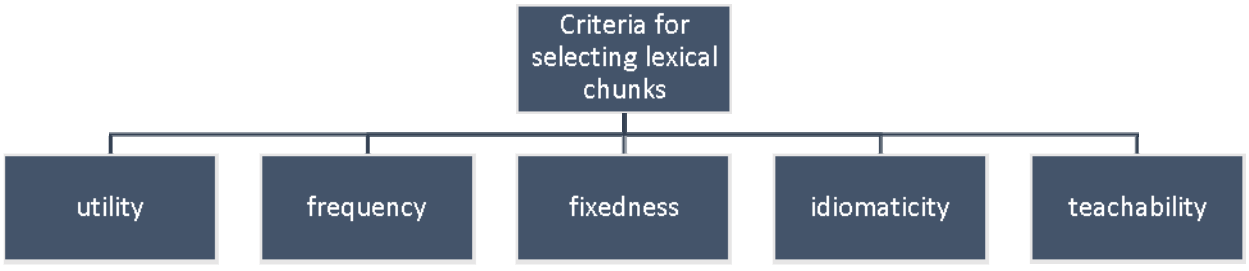
standardised tests such as TOEFL, SAT, etc. Some vocabulary learning applications are integrated with visuals or pictures for learning words. One example of such applications is *vocabulary.com*, where learners can learn and understand myriads of word meanings via learning games and activities. Another example is a lexicon building application, *Personal Lexicon*[1], that allows for creating a language-specific lexicon for learners. In Personal Lexicon, the lexical items, including words, phrases, or expressions, are added and categorised into themes with corresponding translations and language-specific lexical types (nouns, verbs, etc.). The application allows users to take quizzes, track their progress, and share with others. The last example is the *Grammar Girl* website[2] which provides language learning resources and recommends word usages in contexts. However, none of these applications is based on lexical chunks as the starting point or focal point. The reason behind adopting lexical chunks as the starting point in developing a language learning application is because learning language in chunks benefits the learners in terms of facilitating fluent processing and production, conferring indirect or idiomaticity interpretation and providing raw materials for further language development (Thornbury, 2019).

Design Framework

At this point, this research is established on the lexical approach and paradigmatic/syntagmatic relations in second language learning. Moreover, at present, to apply this research to the teaching and learning of second languages, it is undeniable that we need to consider the current pandemic situation in which normal or traditional classroom teaching and learning cannot happen. We, therefore, present the design and development of an educational application called *Bite-sized English* aimed at second language learners of English to learn beyond words by themselves. In designing this application, we aim to answer the question of *which chunks learners should learn and when?* To answer this question, a number of researchers have proposed criteria for selecting lexical chunks for inclusion in language teaching. These criteria (Thonbury, 2019) include:

Figure 2

Criteria for Selecting Lexical Chunks for Language Learners



[1]https://personal-lexicon.com/

[2]https://www.quickanddirtytips.com/grammar-girl

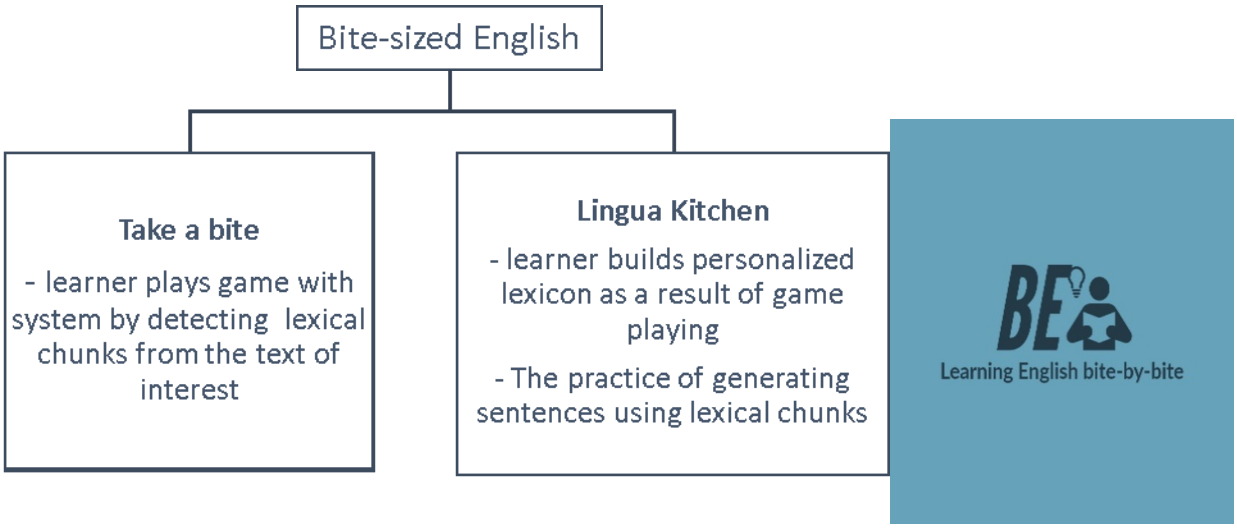
The design of Bite-sized English is based on "utility", meaning that the contents or which chunks to learn are not determined by grammar lessons in textbooks but are selected based on learners' needs, interests and relevance. Therefore, we provide real texts from newspapers, blogs or stories which learners can select and start learning by identifying chunks from such texts in the first component of the system. After that, the application directs learners' attention to the compositional features and usage of chunks in the second component, as discussed in the next section.

Components of the Web Application

Bite-sized English consists of two parts: *Take a Bite,* where users learn lexical chunks from authentic texts, and *Lingua Kitchen*, where learners can build their own lexicon and practice usage in contexts.

Figure 3

Two Components of the "Bite-sized English" Application



The process of the "Take a Bite" component is as follows:

1) Learners select the text of interest from a list of texts, from which they will detect and extract lexical chunks.

2) The learners read the text, and they will be asked to input the lexical chunks that they have found. After that, they can check for chunks that they found against the correct ones from the system.

3) The learners can access more explanations about the lexical chunks as well as give feedback to the system.

The process of "Lingua Kitchen" is described below:

1) The learners can stock their kitchen (i.e. create a vocabulary list or lexicon) by importing the lexical chunks they have learned from the "Take a Bite" part or build from scratch.

2) They can then use the lexical chunks to practice generating sentences and/or adding contexts or pictures that can help them understand and use the chunks better.

The Algorithm for Lexical Chunks Detection: N-gram

To answer the question from the system development point of view on how to detect and extract lexical chunks from texts, we consider using *n-gram*. As mentioned above, in the first component of our system, learners will identify lexical chunks and check or compare them with the answers in our system. Initially, we prepare the set of lexical chunks associated with each text. However, to scale up our system, we need to automatically detect and retrieve lexical chunks from texts, which can be achieved by n-gram. N-grams are sequences of *n* items, where *n* stands for any natural number (1, 2, 3, 4, etc.) of words or linguistic units. For example, the string *detect fake news* contain three 1-grams: *detect, fake, news*, two 2-grams (bigrams): *detect fake* and *fake news* and one 3-gram (trigram): *detect fake news*. N-gram is the basis for the language model, which is the probability distribution of word sequences over sentences from real language use in corpora. Based on n-grams extracted from a large corpus, a language model computes the likelihood of an entire string of *n* items ('joint probability') and/or the likelihood of a single upcoming item given n-1 previous items ('conditional probability'). Estimates of these probabilities generated by an n-gram-based language model are used in a variety of practical applications, including automatic lexical chunks detection, which is the next phase of our research.

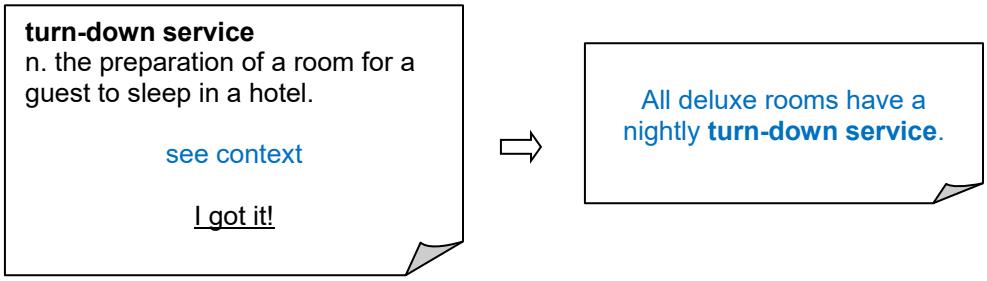
Discussion

In designing "Bite-sized English", we began by considering the assumption about the language production process and language competence as well as the paradigmatic and syntagmatic knowledge. Furthermore, we adopted a lexical approach to language learning especially second language learning. The concept of the lexical chunk is the key of the lexical approach because it is the bridge between word and phrase levels which can be used as building blocks of producing sentences, paragraphs, etc. In learning with the lexical approach in our application, we tried to answer the question of which lexical chunks to learn and when. We based our design framework on "utility" or usefulness and relevance to the learners' purpose in learning. This can be seen from the materials we use, which are natural and authentic texts from newspapers, blogs, online texts and so forth. Learners' awareness of or attention to lexical chunks is raised in using our application, in the "Take a Bite" part, by having learners identify the chunks by themselves from such materials, and this corresponds to Lewis' first recommendation (Lewis, 1993; 1997) in applying the lexical approach to language learning.

After identifying lexical chunks from texts, when learners create a lexicon in the "Lingua Kitchen" part of our application, they learn about the paradigmatic and syntagmatic properties of lexical chunks, as can be seen from an example entry of "turn-down service" lexical chunk below.

Figure 4

Sample Entry for the Lexical Chunk "Turn-down Service" in Lingua Kitchen



The paradigmatic information of "turn-down service" includes the categorical information (i.e. noun) and definition, while the context of usage shows the syntagmatic information of how the chunk is applied and its relation to other elements in the example sentence. Learners will exploit this knowledge in generating sentences using lexical chunks in the practice part of Lingua Kitchen. This allows learners to engage more with the usage of lexical chunks and remember them for further application (Schmitt & Schmitt, 2020).

With the intention to build our design framework based on the previously mentioned linguistic background and assumptions, we have also sought to evaluate *Bite-sized English* as a language learning application. We adopt Roselle-Aguilar's (2017) proposed framework for evaluating language learning applications, which consists of four aspects: *language learning, pedagogy*, *user experience* and *technology*. Regarding the *language learning* aspect, our application concerns vocabulary learning, and it can be evaluated based on specific activities for vocabulary acquisition. In our case, this would be lexical chunks detection and extraction in the "Take a Bite" part. For *pedagogy*, there are three main points based on which our application can be evaluated: *teaching, scaffolding* and *engagement*. Our application is not merely for testing vocabulary knowledge or skills but presents the learners with the materials and allow the users to detect and extract lexical chunks with which a lexicon can be built. Thus, our application has teaching and scaffolding features. Moreover, learners can engage in more challenging tasks by practice generating sentences with lexical chunks. Subsequently, user experience and technology aspects for evaluating language learning applications concern human-computer interaction issues such as interface, navigation and so forth. We take these aspects into consideration for evaluating the beta version of our application later.

Finally, we want to emphasise the significance of having self-learning applications during the time of the COVID-19 pandemic. Since the design of Bite-sized English is aimed at self-learning anytime and anywhere, we integrate the strategies that make students engage and learn better with the system without teachers/supervisors in classroom settings. The building, customising and practising of lexical chunks in our system, in particular in Lingua Kitchen, are based upon the trajectory that supports learning. That is, learners can incrementally practice and apply the new knowledge and skills by using the chunks to generate sentences beyond rote memorisation.

Conclusion

To summarise, in our research, we investigate the issue of second language learning within the linguistic and language production framework. We explain language learning and competence by knowledge of paradigmatic and syntagmatic relations as well as lexical approach. With this basis, we have designed an educational web application for second language learning, *Bite-sized English*, that incorporates the learning of words and phrases hands-in-hands with grammar.

Bite-sized English consists of two components: Take a bite and Lingua Kitchen. In the former, learners can detect lexical chunks from authentic texts and check for correctness. In the latter, learners can build a lexicon based on the lexical chunks they previously learned as well as practice usages of lexical chunks in generating sentences. The design of our application is assessed based on the evaluation framework for language learning applications, including language learning, pedagogy, user experience and technology. As a result of a preliminary evaluation, our application is equipped with pedagogical activities and features that allow for engaging vocabulary learning and beyond.

References

Biber, D. (1999). *Longman grammar of spoken and written English*. Longman.

Culler, J. D. (1976). *Ferdinand de Saussure* (7th ed.). Cornell University Press.

Cutler, A. (Ed.). (1982). *Slips of the tongue and language production.* De Gruyter Mouton. https://doi.org/10.1515/9783110828306

Dell G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283-321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., Kelley, A. C., Hwang, S., & Bian, Y. (2021). The adaptable speaker: A theory of implicit learning in language production. *Psychological Review*, 128(3), 446–487. https://doi.org/10.1037/rev0000275

Fromkin, V. (Ed.) (1973). *Speech errors as linguistic evidence*. Mouton.

Fromkin, V. (Ed.) (1980). *Errors in linguistic performance: slips of the tongue, ear, pen, and hand*. Academic Press.

Gordon, J. & Dell G. S. (2003). Learning to divide the labour: An account of deficits in light and heavy verb production. *Cognitive Science,* 27(1), 1-40. https://doi.org/10.1016/S0364-0213(02)00111-8

Jaeger, J. (1992). Phonetic features in young Children's slips of the tongue. *Language and Speech,* 35(1-2), 189-205. https://doi.org/10.1177%2F002383099203500215

Krashen, S. (1993). *The power of reading: Insights from the research*. Libraries Unlimited.

Levelt, W. (1989). *Speaking: From intention to articulation.* MIT Press.

Levelt W. J., Roelofs A., & Meyer A. S. (1999). A theory of lexical access in speech production. *Behaviour and Brain Sciences,* 22(1), 1-38; discussion 38-75. https://doi.org/10.1017/s0140525x99001776.

Lewis, M. (1993). *The lexical approach*. Language Teaching Publications.

Lewis, M. (1997). *Implementing the lexical approach*. Language Teaching Publications.

Limacher, U. (2017). *Language acquisition vs. language learning*. http://utesinternationallounge.com/language-acquisition-versus-language-learning/

Paradis, M. (1998). Chapter 29 - Language and communication in multilinguals. In B. Stemmer & H. A. Whitaker (Eds.), *Handbook of Neurolinguistics* (pp.417-430), Academic Press, https://doi.org/10.1016/B978-012666055-5/50033-2

Rosell-Aguilar, Fernando (2017). State of the app: A taxonomy and framework for evaluating language learning mobile applications. *CALICO Journal*, 34(2), 243–258. https://doi.org/10.1558/cj.27623

Saussure, F. D. (1959). *Course in general linguistics*. Philosophical Library.

Schmitt, N. & Schmitt, D. (2020). *Vocabulary in language teaching* (2nd ed.). Cambridge University Press.

Stemberger, J. P., Pisoni, D. B., & Hathaway, S. N. (1985). Effects of alcohol intoxication on phonological errors in normal speech. *Research on Speech Perception*, Progress report, 11 (pp. 95-404). Speech Research Laboratory, Department of Psychology, Indiana University.

Thornbury, S. (2019). *Learning language in chuncks*. Cambridge University Press. https://www.cambridge.org/elt/blog/wp-content/uploads/2019/10/Learning-Language-in-Chunks.pdf

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press. https://doi.org/10.1017/CBO9780511519772

Acknowledgements

This work has been supported by the research grant from the Faculty of Engineering, King Mongkut's University of Technology Thonburi to Asst. Prof. Dr. Nuttanart M. Facundes.

Funding

Not applicable.

Ethics Declarations

Competing Interests

No, there are no conflicting interests.