LEARN Journal: Language Education and Acquisition Research Network
ISSN: 2630-0672 (Print) | ISSN: 2672-9431 (Online)
Volume: 18, No: 1, January – June 2025

Language Institute, Thammasat University
https://so04.tci-thaijo.org/index.php/LEARN/index

The Role of Input Frequency and Different Proficiency Levels on the Perception of English Nominal Suffixes by L1 Thai Learners: A Case of the Usage-Based Account

Chanikarn Thatchatham[a], Nattama Pongpairoj[b,*]

[a] chanikarn.nu@kmitl.ac.th, English as an International Language, Graduate School, Chulalongkorn University, Thailand
[b] pnattama@chula.ac.th, English as an International Language, Graduate School, Chulalongkorn University, Thailand
[*] Corresponding author, pnattama@chula.ac.th

APA Citation:
Thatchatham, C., & Pongpairoj, N. (2025). The role of input frequency and different proficiency levels on the perception of English nominal suffixes by l1 Thai learners: A case of the usage-based account. *LEARN Journal: Language Education and Acquisition Research Network. 18*(1), 883-913. https://doi.org/10.70730/VDET8305

Received 02/07/2024
Received in revised form 17/12/2024
Accepted 24/12/2024

ABSTRACT

This study investigated how input frequency (i.e., type frequency and token frequency) and proficiency levels enhanced the perception of English nominal suffixes by first language (L1) Thai learners. Based on the Usage-based Account (Tomasello, 2003), it was hypothesized that input frequency, i.e., token frequency (frequency of derived forms containing the particular suffix) and type frequency (suffix frequency), facilitates SLA. A Grammaticality Judgement Test (GJT) was administered to 60 L1 Thai learners at the intermediate and the advanced proficiency levels, 30 per group. The four frequency conditions were as follows: Condition 1 (HH) – high type and high token frequency (e.g., 'alteration'); Condition 2 (HL) – high type and low token frequency (e.g., 'chemist'); Condition 3 (LH) – low type and high token frequency (e.g., 'dependence'); and Condition 4 (LL) – low type and low token frequency (e.g., 'partnership'). The results

showed significant effects of input frequency and proficiency levels. However, the interaction between them was not significant. The study also revealed that the intermediate group perceived Condition 3 (LH) most accurately, whereas Condition 1 (HH) was perceived most accurately by the advanced group. The findings supported the Usage-based Account, indicating that input frequency influenced SLA of English nominal suffixes, with token frequency having a greater impact than type frequency.

Keywords: input frequency, type frequency, token frequency, English nominal suffixes, Usage-based Account

Introduction

One area of English derivational morphemes that is likely to pose challenges to L1 Thai learners is English nominal suffixes. This is because nominalization, which is the process of deriving a noun from a word of other syntactic classes (Prasithrathsint, 1996), differs significantly in English and Thai.

In English, nouns are derived by adding nominal suffixes, known as nominalizers, to several types of roots including verbs, adjectives, and nouns (Hamawand, 2011). For instance, the suffix '-ness' derives 'happy' into 'happiness', and '-ship' derives 'friend' into 'friendship'. English nominalization allows a wide variety of nominal suffixes, which can attach to roots of different syntactic categories. In contrast, while Thai also forms derived nouns through the process of derivation, it exhibits key differences in the mechanisms involved in the number and types of nominal affixes and the types of roots. In Thai, nouns are predominantly created by adding one of the two nominal prefixes, 'kaan' or 'khwaam', to verbal roots. These prefixes differ in their distribution and the meanings they convey from the original verbs (Prasithrathsint, 2005). Unlike English, Thai does not have a diverse range of nominal affixes or the ability to derive nouns from adjectival or nominal roots, further distinguishing its approach to nominalization from that of English.

Second language (L2) English learners' knowledge of English nominal suffixes could be promoted and developed through input frequency (Sayer & Abdulsalam, 2018). Input frequency, which falls under the Usage-based Theory, is considered a significant determinant that can enhance L2 English learners' acquisition of English nominal suffixes. This is because the frequency of exposure to language helps 'entrench' – or strengthen the abstract representation or the mental schema of particular linguistic patterns

(Dąbrowska, 2004). Input frequency could be classified into token frequency and type frequency (Bybee & Beckner, 2010; Ellis & Collins, 2009).

Token frequency refers to the frequency count of a particular word or phrase in natural language use (Bybee & Beckner, 2010). According to Croft and Cruse (2004), token frequency is the vital determinant for storing the whole form. Thus, words or phrases with high token frequency are entrenched much more easily (Langacker, 1987, as cited in Croft & Cruse, 2004). As far as type frequency is concerned, it refers to the frequency counts of particular linguistic items (e.g., morphemes, phonemes, and words) that can be replaced in the given slot in the pattern (Bybee & Beckner, 2010). However, the exact role of input frequency in SLA remains incompletely understood (Almulla, 2015). Croft and Cruse (2004) pointed out that it is difficult to identify the role of type frequency and token frequency in the entrenchment of linguistic patterns. Thus, the role of input frequency, together with the exact effects of token frequency and type frequency, is still an ongoing issue in SLA and requires further investigation.

There have been several studies on the acquisition of English derivational morphemes by learners from different L1 backgrounds, including Kuwaiti (Alotaibi & Alotaibi, 2017) and Kurdish (Sayer & Abdulsalam, 2018). As long as research in the Thai context is concerned, there was only one study (Chiarakiat, 2019) that examined perception of English adjectival suffixes. Of all the studies mentioned earlier, only two studies (Sayer & Abdulsalam, 2018; Chiarakiat, 2019) employed input frequency as the independent variable. To the best of our knowledge, there have not been any studies investigating the perception of English nominal suffixes in the Thai context. Thus, this study filled the gap by exploring the role of input frequency in perception of English nominal suffixes by L1 Thai learners of different proficiency levels, i.e., intermediate and advanced. Three research questions were formulated as follows.

1. To what extent will input frequency enhance the perception of English nominal suffixes by intermediate and advanced proficiency levels?
2. How do two types of input frequency, i.e., token frequency and type frequency, have an impact on perception of English nominal suffixes by L1 Thai learners of intermediate and advanced proficiency levels?
3. How does input frequency have an interaction with proficiency levels, i.e., intermediate and advanced proficiency levels, by L1 Thai learners?

Literature Review

The Usage-Based Theory

The Usage-based Theory emphasizes that linguistic structures are shaped by use rather than being innate (Croft & Cruse, 2004; Dąbrowska, 2004). Zyzik (2009) further proposed that language acquisition is driven by input from usage, with learners utilizing their cognitive abilities to process this input into abstract mental representations.

The Apparatus in the Usage-Based Theory: Exemplars and Networks

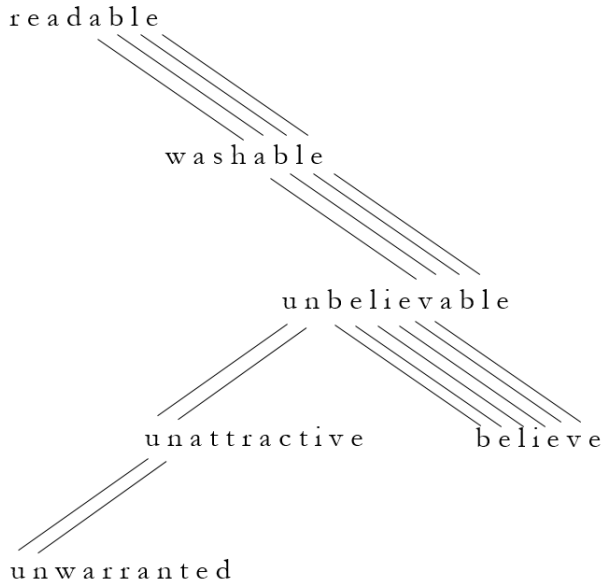
It was proposed under the Usage-based Theory that speakers have "rich memory representations" that store their experiences with language. (Goldinger, 1996, as cited in Bybee & Beckner, 2010, p.833). This mental representation of linguistic input can be termed as a representational apparatus, which could be categorized into two models: exemplars and networks.

The first apparatus is exemplars, or mental categories formulated from experienced linguistic inputs that have shared features (Pierrehumbert, 2001). They range in size, from individual linguistic units like consonants to paragraphs (Bybee, 2013). Exemplars are created through the process of categorization, where similar items are grouped together (Bybee, 2013). When an input token matches an already existing exemplar, it is immediately assigned to that exemplar, which strengthens the exemplar (Bybee & Beckner, 2010). For instance, nouns derived with the suffix '-ship' (e.g., 'friendship', 'relationship', 'apprenticeship') are likely mapped to the exemplar of the nominal suffix '-ship'. This process leads to the formation of 'exemplar clouds' or constituent categories for the nominal suffix '-ship' (Bybee, 2013).

The other apparatus is networks, which refer to the mental representations of common linguistic features at various levels (Bybee & Beckner, 2010). In other words, networks consist of groups of exemplars that share mutual similarities. When linguistic inputs share common features with exemplars or exhibit slight differences, their shared characteristics are located nearby in the representation, thereby creating a network (Bybee & Beckner, 2010).

Figure 1

The network of 'unbelievable' and connections to related words (Bybee & Beckner, 2010, p.835)



In Figure 1, the center of this network is the derived word 'unbelievable' with a mental schema of [stem[1]+ '-able']. It is a member of exemplar clouds containing the suffix '-able' (e.g., 'readable' and 'washable'). Simultaneously, 'unbelievable' falls within the category of words sharing the root 'believe'. Furthermore, it is a member of the mental schema ['un-' +stem], which includes other derived words like 'unattractiveness' and 'unwarranted'. Hence, there are several exemplars including the exemplar of the words attached by the suffix '-able', words containing 'believe', and words attached by the prefixes 'un-' in these networks.

Input Frequency

Input frequency is a major issue of the Usage-based Theory as the frequent recurrence of a specific structure in running texts can significantly impact language acquisition. Cognitive advocates like Bybee and Beckner (2010), Croft and Cruse (2004), and Dąbrowska (2004) concurred that input frequency helps reinforcing exemplars and strengthening the mental

representation. Input frequency is divided into two types: token frequency and type frequency (Bybee & Beckner, 2010; Ellis & Collins, 2009).

Token frequency is the frequency counts of the derived word types containing the particular affix (Laws & Ryder, 2014). In this study, it represents the total number of occurrences of all derived nouns containing the specific nominal suffix in the MorphoQuantics Corpus (https://morphoquantics.co.uk). For example, the nominal suffix '-hood' has a token frequency of 318, which means that the combined frequency of all derived nouns containing '-hood' (e.g., 'adulthood', 'boyhood', and 'brotherhood') is 318. Token frequency helps determine the storage of word forms (Croft & Cruse, 2004). That is, the more often a linguistic form is produced, the more robust a schema is, resulting in "its ultimate storage as a conventional grammatical unit" (Croft & Cruse, 2004, p.292). Bybee and Beckner (2010) proposed two effects from learners' frequent exposure to high token frequency linguistic items: the reduction and conserving effects. The reduction effect refers to the phenomenon where linguistic items with high frequency tend to undergo more reduction compared to low-frequency ones. It could also occur in several linguistic areas including phonetics, and syntax (Bybee & Thompson, 1997). For example, when the phrase '(be) supposed to' undergoes this process, it reduces to [spostə] in production. In this case, the infinitive 'to' merges with the verb, becoming chunk sequences due to high frequency. The second effect of token frequency is the conserving effect, which pertains to the resistance of complex forms to reformulation (Bybee & Beckner, 2010). This preservation promotes fluency and enhances memory representations, leading to quicker lexical access. For example, the derived adjective 'visible', which is of high token frequency, tends to be conserved as a whole ('visible') rather than parsed into derivational units ('vis-' + '-able').

Type frequency refers to the count of lexical items that can occupy a specific position in a structure (Ellis & Collins, 2009). In this study, it refers to the frequency of target English nominal suffixes. For example, the nominal suffix '-or' has a type frequency of 224, indicating it appears in 224 different words. It is widely acknowledged that type frequency enhances the productivity of a structure (Bybee & Beckner, 2010). That is, patterns or structures applicable to a wide range of items tend to be applied to new items as well (Bybee & Beckner, 2010). For example, the nominal suffix '-ness', which is of high type frequency based on the MorphoQuantics corpus, is likely to be applicable to a variety of adjectives (e.g., 'happy', 'joyful', and 'loud'), forming derived nouns (e.g., 'happiness', 'joyfulness', and 'loudness'). Thus, the nominal suffix '-ness' is of high productivity. Type frequency also enhances the ability to parse of the construction (Hay & Baayen, 2003). For example, according to Bybee and Beckner (2010), when learners encounter

the word 'happiness' without prior knowledge of related words, they may struggle to infer that 'happiness' is composed of two morphemes. However, if they experience the nominal suffix '-ness' that is attached to other adjectives (e.g., 'joyfulness', 'sadness' and 'sickness'), they would be aware that '-ness' is a suffix that can be attached to adjectives. So, a degree of type frequency of a particular linguistic structure must be adequately high so that the learner could develop the ability to parse the structure (Bybee & Beckner, 2010), promoting the readiness for applying for a novel use.

Croft and Cruse (2004) proposed that token frequency and type frequency jointly contribute to forming mental representations of linguistic forms. When both token frequency and type frequency are high, token frequency reinforces the entrenchment of the form, while type frequency enhances understanding of its productivity. As productivity increases, learners develop the ability to analyze and apply new linguistic elements to similar constructions (Ellis & Collins, 2009). Therefore, token frequency and type frequency have combined effects on the acquisition of linguistic constructions.

Nominalization in English and Thai

Nominalization is the process of deriving a noun from a word of another class (Prasithrathsint, 1996). It varies between English and Thai. English has several nominal suffixes, whereas Thai has only two.

Nominalization and Nominalizers in English

In English, nouns are formed using nominal suffixes or nominalizers (Hamawand, 2011). These suffixes are bound morphemes attached to free morphemes like verbs, adjectives, or nouns.

First, nominalizers attached to verbal roots fall into two types: simple nouns or agent nouns (Hamawand, 2011). Simple nouns denote things, actions, or abstract concepts, while agent nouns describe individuals performing specific actions. For example, suffixes such as '-al', '-ion', and '-ment' create simple nouns ('withdrawal', 'communication', and 'payment', respectively). Agent nouns like 'writer' and 'employee' are formed with suffixes such as '-er' and '-ee'.

Second, nominalizers attached to adjectival roots include suffixes like '-cy', '-ity', and '-ness' and form simple nouns (Hamawand, 2011). For instance, '-cy', '-ity', and '-ness' attached to adjectives like 'fluent', 'agile', and 'calm' derive nouns like 'fluency', 'agility', and 'calmness'.

The third type of English nominal suffixes includes those attached to nominal roots, known as class-maintaining suffixes (Hamawand, 2011). These

suffixes retain the word class of the stem they attach to. For example, '-dom', '-hood', and '-ship' attached to 'king', 'child', and 'friend' form 'kingdom', 'childhood', and 'friendship', respectively.

Nominalization and Nominalizers in Thai

In Thai, nouns can be derived through a process called 'lexical nominalization,' where abstract nouns are formed by adding nominalizers (Prasithrathsint, 2005). There are two primary nominalizers: 'kaan' and 'khwaam' (Prasithrathsint, 2005).

The nominalizer 'kaan-' derives from the noun 'kaan', meaning 'affair' (Prasithrathsint, 1997). As a prefix, it conveys meanings such as 'matters of ...' and 'act of ...' (Smyth, 2002, p.29). Conversely, 'khwaam' originally denotes 'a sense of a matter' (Prasithrathsint, 1997).

Prasithrathsint (2005) provided the criteria for the selection between 'kaan' and 'khwaam' based on the meanings of verbs including perceptible verbs, imperceptible verbs, and balanced verbs.

First, perceptible verbs refer to action or non-action verbs whose meanings are clear to native Thai speakers (Prasithrathsint, 2005). These verbs permit only the attachment of 'kaan'. For example:

(1) càtkaan (to manage) – kaan-càtkaan (management)
(Prasithrathsint, 2005, p.76)

Secondly, imperceptible verbs describe characteristics or qualities of a person or thing, resembling adjectives in English. The nominalizer 'khwaam' is used with this type of verb. For example:

(2) sàdùak (convenient) – khwaam-sàdùak (convenience)
(Prasithrathsint, 2005, p.77)

Thirdly, balanced verbs exhibit characteristics of both perceptible and imperceptible verbs. Both 'kaan' and 'khwaam' can be attached to balanced verbs, each conveying distinct inherent meanings. For example:

(3) rák (to love) – kaan-rák (loving)
– khwaam-rák (love)
(Prasithrathsint, 2005, p.78)

Previous Studies

To date, there have been only two studies (Sayer & Abdulsalam, 2018; Chiarakiat, 2019) that examined the role of input frequency on the acquisition of English derivational morphemes by L2 learners.

Sayer and Abdulsalam (2018) investigated how L1 Kurdish university students comprehended and produced English derivational morphemes. There were 112 third and fourth-year English-major students from the

University of Human Development, Iraq, categorized into four proficiency levels: pass, medium, good, and very good. A multiple-choice test was employed to assess comprehension and a Fill-in-the-Blank Test was utilized to evaluate production skills regarding twelve English derivational suffixes. These suffixes were grouped based on type frequency in the MorphoQuantics corpus including high (e.g., '-ly', '-ion'), medium (e.g., '-ous', '-or'), and low (e.g., '-ways', '-ie'). The findings indicated that higher proficiency levels correlated with better performance in both tasks. It was found that input frequency had a significant effect only on comprehension. The findings aligned with Schmitt and Zimmerman's (2002) assertion that perceiving derivatives is simpler than producing them.

Chiarakiat (2019) investigated the role of input frequency on the perception of English adjectival suffixes by L1 Thai university students. Thirty first-year university students were recruited to perform in the Grammaticality Judgement Test (GJT). The target items were divided into four conditions based on token frequency and type frequency information from the MorphoQuantics corpus: Condition 1 – HIGH type and HIGH token frequency (HH) (e.g., 'environmental'), Condition 2 – HIGH type and LOW token frequency (HL) (e.g., 'dusty'), Condition 3 – LOW type and HIGH token frequency (LH) (e.g., 'regular'), and Condition 4 – LOW type and LOW token frequency (LL) (e.g., 'plausible'). The results revealed that the input frequency had a significant main effect on the perception of English adjectival suffixes, with stronger influence from token frequency.

To the best of our knowledge, there were no previous studies investigating how input frequency together with proficiency levels influence the perception of English nominal suffixes. Thus, this study aimed to explore the role of input frequency and proficiency levels on the perception of English nominal suffixes.

Research Methodology

Participants

There were three participant groups: two groups of L1 Thai speakers and one group of native English speakers. The first two groups consisted of first-year undergraduate students at Chulalongkorn University in Bangkok, Thailand. They were sorted into two proficiency levels: intermediate (B1) and advanced (C1), determined by CU-TEP[2] scores. A score between 35 and 69 corresponds to CEFR level B1 (Wudthayagorn, 2018), while a score between 99 and 120 corresponds to CEFR level C1 (Wudthayagorn, 2018). The L1 Thai participants were non-English majors from various faculties, studying in the 'Experiential English I' course offered by the Chulalongkorn University

Language Institute. Their average age was 18.43 (*SD* = 0.84). All the participants were native Thai speakers who completed compulsory education and had studied English for 12 years in schools where Thai was the primary language of instruction. They were not enrolled in an English Program (EP), Intensive English Program (IEP), or Bilingual Program, nor were they exchange students or individuals who had lived in an English-speaking country for more than three consecutive months. The third group comprised ten native English speakers. The data obtained from the native speaking group was used as the baseline data.

The Selection of Data

This study focused on noun-forming suffixes from verbal, adjectival, and nominal roots. Target nominal suffixes were selected based on the type and token frequency data available on the MorphoQuantics corpus (https://morphoquantics.co.uk), in which the data on type and token frequency of derivational morphemes could be found. In this study, type frequency refers to the frequency counts of targeted English nominal suffixes (Laws & Ryder, 2014). For instance, the nominal suffix '-or' demonstrates a type frequency of 224, indicating that it appears in 224 distinct words. Token frequency refers to the frequency counts of the derived word types containing the particular affix from the spoken component of the British National Corpus (BNC) (Laws & Ryder, 2014). For instance, the nominal suffix '-ence' demonstrates a token frequency of 4,075, which represents the total occurrences of all derived nouns containing this suffix. The classification of nominal suffixes into high and low type and token frequency groups was based on distribution patterns and clear cutoff points. For type frequency, suffixes with counts above 100 (e.g., '-ation' (267[3]) and '-ism' (189)) were categorized as high, while those below 100 (e.g., '-ence' (101) and '-ship' (54)) were low. For token frequency, suffixes exceeding 10,000 (e.g., '-ation' (12,7503[4])) were high, while those below 10,000 (e.g., '-ency' (1,481)) were low. There were 8 English nominal suffixes chosen in this study. Four of them, i.e., '-ation' (267), '-or' (224), '-ist' (210) and '-ism' (189), were sorted into the high type frequency group, while the other four of them, i.e., '-ence' (101), '-ure' (71), '-ship' (54), and '-ency' (34), were of low type frequency. Regarding token frequency, four English nominal suffixes were of high token frequency '-ation' (12,750[3]), '-or' (10,814), '-ence' (9,577), and '-ure' (11,113), while the other four nominal suffixes: '-ist' (2,423), '-ism' (1,748), '-ship' (2,366), and '-ency' (1,481) were of low token frequency. An independent samples t-test was conducted to see whether there are significant differences in 1) the mean frequency counts of nominal suffixes with high type frequency versus low type frequency, and 2) the mean frequency counts of nominal

suffixes with high token frequency versus low token frequency. The result revealed that the nominal suffixes with high type frequency (M = 222.5, SD = 32.97) had higher type frequency counts than those with low type frequency (M = 65, SD = 28.37), demonstrating a significant difference, t(6) = 7.242, p < .001. At the same time, the nominal suffixes with high token frequency (M = 11018.5, SD = 1229.61) had higher token frequency counts than those with low token frequency (M = 2004.5, SD = 463.92), showing a significant difference, t(6) =13.718, p < .001.

In this study, there were four conditions of nominal suffixes including Condition 1: HIGH type and HIGH token frequency (HH), Condition 2: HIGH type and LOW token (HL), Condition 3: LOW type and HIGH token (LH), and Condition 4: LOW type and LOW token frequency (LL) (see Table 1).

Table 1

Type Frequency and Token Frequency of Each Condition

Conditions	Nominal Suffixes	Type frequency	Token frequency
Condition 1 (HH)	-ation	267	12570
	-or	224	10814
Condition 2 (HL)	-ist	210	2423
	-ism	189	1748
Condition 3 (LH)	-ence	101	9577
	-ure	71	11113
Condition 4 (LL)	-ship	54	2366
	-ency	34	1481

The Data Elicitation Task

Grammaticality Judgement Test

The Grammaticality Judgement Test (GJT) was set to measure how the intermediate and advancedproficiency participants perceived English nominal suffixes with high and low type frequency. There were 40 sentences, comprising 16 target test items and 24 distractors (see Appendix A). There were four target test items in each condition. All the sentences were derived from the concordance lines available on the Corpus of Contemporary American English (COCA) and simplified to ensure the participants' comprehension. The test was validated through the use of the Index of Item-objective Congruence (IOC) by three experts who were native English instructors. The variables were controlled as follows. All derived nouns in the

test sentences must have no premodifiers such as determiners (quantifiers, demonstrative determiners, and possessive determiners), adjectives, and adverbs. The only premodifiers permitted were articles ('a', 'an', and 'the'). Conversely, the distractors were derived words from other parts of speech, such as adjectives (e.g., 'fruitful' and 'commercial') and verbs (e.g., 'idealize' and 'worsen').

Table 2

The List of Derived Nouns in the GJT Test

Conditions	Nominal suffixes	Type frequency	Token frequency	Derived nouns (word frequency)
1 (HH)	-ation	267	12570	alteration (107)
High type and				combination (133)
High token	-or	224	10814	contractor (142)
				moderator (127)
2 (HL)	-ist	210	2423	chemist (99)
High type and				journalist (90)
Low token	-ism	189	1748	criticism (126)
				mechanist (167)
3 (LH)	-ence	101	9577	correspondence (97)
Low type and				dependence (118)
High token	-ure	71	11113	closure (174)
				agriculture (149)
4 (LL)	-ship	54	2366	partnership (183)
Low type and				apprenticeship (61)
Low token	-ency	34	1481	contingency (92)
				efficiency (136)

In each condition, there were two target test items that were grammatically correct and two that were not, requiring the test takers to correct them. Examples of the target test items were shown below.

1. 1a. I think the <u>alteration</u> is important to the contract. ( ) _____
   1b. The moderatist kept the discussion on track during the meeting. ( ) _____

Data Collection

The data was collected online via Google Forms. Online data collection allowed the participants flexibility to choose convenient time slots, ensuring minimal disruption to their class time while maintaining the integrity of the study's data collection process. The task took 30 minutes, and each participant was also observed via the Zoom program to ensure that they completed the GJT within the time limitation.

Data Analysis

There were two variables in this study: input frequency and proficiency levels. The data from the GJT were analyzed as follows:

First, descriptive statistics were conducted to find the mean, percentage, and standard deviation. Second, a one-way ANOVA was performed to see the effect of input frequency on the GJT scores to determine whether input frequency had a significant main effect in each proficiency group or not. Third, as the effect of input frequency was significant, a subsequent post hoc test, Tukey's HSD, was employed to identify significant differences among the means of each condition of input frequency. Fourth, a 2 × 4 two-way mixed ANOVA was performed to explore the impact of proficiency levels and input frequency, together with the interaction between them.

The data from the native English-speaking group were used as the baseline data to compare with those from the non-native groups.

Results and Discussion from the GJT

This section is divided into two subsections. The first subsection reports the GJT results from the two L1 Thai groups i.e., intermediate and advanced groups, together with the native English-speaking group, followed by the discussion on the role of input frequency (Research Question 1) and the role of type and token frequency (Research Question 2). The second subsection provides the comparison GJT results from the two proficiency groups and discusses the role of input frequency and proficiency levels (Research Question 3).

Results from the GJT and discussion on the role of input frequency on the perception of English nominal suffixes

Results from the GJT from each group

Table 3

Results on the Correct Answers of the GJT from the Intermediate L1 Thai Group

GJT Scores of the intermediate group (30 participants)					
Condition	**Rank**	**Total Scores (120)**	**Percentages**	**Mean**	***SD***
Condition 1 (HH)	2	85	70.83	2.83	0.70

Condition 2 (HL)	3	75	62.50	2.50	0.97
Condition 3 (LH)	1	87	72.50	2.90	0.71
Condition 4 (LL)	4	72	60	2.40	0.62

Table 4

Results on the Correct Answers of The GJT from the Intermediate L1 Thai Group Based on the Different Types of Roots

Types of roots	Number of derived words[5]	Full scores (number of derived words × the number of participants in each proficiency group)	Scores	Percentages
Verbal roots	7	210	148	70.48
Nominal roots	5	150	84	56
Adjectival roots	2	53	43	71.67
Bound base[6]	2	56	44	73.33

The intermediate group's perception rates of derived words from bound bases, adjectival roots, and verbal roots, were at approximate levels, i.e., 73.33%, 71.67%, and 70.83%, respectively. Their perception rate for nominal roots was much lower, i.e., 56%.

Table 5

Results on the Correct Answers of the GJT from the Advanced L1 Thai Group

GJT Scores of the advanced group (30 participants)					
Condition	**Rank**	**Total Scores (120)**	**Percentages**	**Mean**	***SD***
Condition 1 (HH)	1	109	90.83	3.63	0.49
Condition 2 (HL)	2	102	85	3.40	0.67
Condition 3 (LH)	3	100	83.33	3.33	0.66
Condition 4 (LL)	4	92	76.67	3.07	0.69

The patterns of judgements for the target test items differed slightly in both proficiency groups. The intermediate group performed best in Condition 3 (LH) (M = 2.90), followed by Condition 1 (HH) (M = 2.83), and Condition 2 (HL) (M = 2.50). However, the advanced group received the highest mean in Condition 1 (HH) (M = 3.63), followed by Condition 2 (HL) (M = 3.40), and Condition 3 (LH) (M = 3.33). Condition 4 (LL) received the lowest means from both proficiency groups (intermediate: M = 2.40,

advanced: $M = 3.07$). That is, the intermediate group were most sensitive to those with low type and high token frequency, while the advanced group were most sensitive to the target test items with high type and high token frequency. Both proficiency groups perceived the target test items with low type frequency and low token frequency the least accurately.

Table 6

Results on the Correct Answers of the GJT from the Intermediate L1 Thai Group based on the Different Types of Roots

Types of roots	Number of derived words	Full scores (number of derived words × the number of participants in each proficiency group)	Scores	Percentages
Verbal roots	7	210	177	84.29
Nominal roots	5	150	110	73.33
Adjectival roots	2	53	53	88.33
Bound base	2	56	56	93.33

The intermediate group judged the target test items containing bound bases at the highest rate (93.33%), followed by adjectival roots (88.33%), and verbal roots (84.29%). Their perception of nominal roots was at the lowest rate, i.e., 56%.

Table 7

Results of the One-Way ANOVA Analysis for the GJT from the Advanced L1 Thai Group

Source	Sum of Square	*df*	Mean Square	*F*	*p-value*
Between Groups	5.425	3	1.808	3.105	.029*
Within Groups	67.567	116	.582		
Total	72.992	119			

Significance level: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

In the intermediate group, the one-way ANOVA revealed a significant effect for input frequency at the $p < .05$ level ($F(3, 116) = 3.105$, $p < .05$).

As the effect of the input frequency was significant, a post hoc, using Tukey's HSD, was performed to identify significant differences among the means of each condition from the intermediate L1 Thai group.

Table 8

Results on the Comparisons of Means from Each Condition in the GJT by the Intermediate L1 Thai Group from the Post Hoc Test

Conditions	*n*	Mean	*SD*	Tukey's HSD Comparisons			
				1 (HH)	2 (HL)	3 (LH)	4 (LL)
1 (HH)	30	2.83	.69		.333	.987	.130
2 (HL)	30	2.50	.97	.333		.183	.957
3 (LH)	30	2.90	.71	.987	.183		.059
4 (LL)	30	2.40	.62	.130	.957	.059	

Significance level: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Post hoc testing with Tukey's HSD revealed that there were no significant differences among all pairs of means, with $p > .05$ in all the comparisons. That is, Condition 3 (LH) had the highest mean, with no significant differences with any conditions.

Table 9

Results of the One-Way ANOVA Analysis for the GJT from the Advanced L1 Thai Group

Source	**Sum of Square**	***df***	**Mean Square**	***F***	***p-value***
Between Groups	4.892	3	1.631	4.0650	.009**
Within Groups	46.700	116	.403		
Total	51.592	119			

Significance level: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

The one-way ANOVA revealed a significant effect of input frequency at the $p < .01$ level ($F(3, 116) = 4.065$, $p < .01$).

A subsequent post hoc test, Tukey's HSD, was employed to identify significant differences among the means of each condition of input frequency in the advanced L1 Thai group. The results are presented in Table 10.

Table 10

Results on the Comparisons of Means From Each Condition in the GJT by the Advanced L1 Thai Group from the Post Hoc Test using the Tukey's HSD Comparisons

				Tukey's HSD Comparisons			
Conditions	*n*	Mean	*SD*	1 (HH)	2 (HL)	3 (LH)	4 (LL)
1 (HH)	30	3.63	.49		.487	.264	.004**
2 (HL)	30	3.40	.67	.487		.977	.181
3 (LH)	30	3.33	.66	.264	.977		.367
4 (LL)	30	3.07	.69	.004**	.181	.367	

Significance level: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

For the advanced L1 Thai group, post hoc testing with Tukey's HSD revealed significant differences between Condition 1 (HH) and Condition 4 (LL) ($p < .01$). However, no significant differences were observed among the other pairs of means. That is, the mean of Condition 1 (HH) was the highest, significantly surpassing Condition 4 (LL) and showing no significant difference compared to Conditions 2 (HL) and 3 (LH).

Table 11

Results on the Correct Answers of the GJT from the Native-English Speaking Group

GJT Scores of the native English-speaking group (n=10)				
Conditions	**Total Scores (40)**	**Percentage**	**Mean**	***SD***
Condition 1 (HH)	40	100	4	0
Condition 2 (HL)	40	100	4	0
Condition 3 (LH)	40	100	4	0
Condition 4 (LL)	40	100	4	0

The ten native English speakers judged the target test items correctly in all conditions (100%, $M = 4$). This means that they perceived all the target test items, regardless of type frequency and token frequency.

Discussion of the GJT Results on the Role of Input Frequency Together with Type Frequency and Token Frequency

The Role of Input Frequency on the Grammaticality Judgement of English Nominal Suffixes by the L1 Thai Learners and the Native English Group

The first hypothesis stated that input frequency played a role in the perception of English nominal suffixes by the L1 Thai learners. The one-way ANOVA analysis was utilized to compare the means of the four conditions

in the GJT within each proficiency group. The results from the GJT among the intermediate and advanced L1 Thai groups confirmed this hypothesis, demonstrating a significant effect of input frequency on the grammaticality judgement scores in both proficiency groups.

It was found that the intermediate group judged the target test items with low type and high token frequency (Condition 3 (LH) (i.e., '-ence' and '-ure')) most accurately, while the advanced group's most accurate judgement was in Condition 1 (HH), which had English nominal suffixes with high type and high token frequency (i.e., '-ation' and '-or'). Both groups performed least accurately on the target test items with low type and low input frequency (Condition 4 (LL)). This result aligned with Usage-Based Theory (Tomasello, 2003), suggesting that frequent exposure to linguistic patterns facilitates acquisition and storage in mental representations (Ellis & Collins, 2009), resulting in stronger recognition (Bybee & Beckner, 2010). Thus, linguistic forms with high frequency are processed more correctly and faster than those with low frequency (Bertram et al., 2000). This can be seen from the GJT results as both proficiency groups were most sensitive to the target test items with high frequency, suggesting the entrenchment of the target derived nouns with high input frequency in their mental storage. Hence, both proficiency groups were able to access and retrieve the nominal suffixes with frequency occurrences more effectively. In addition, input frequency played a crucial role in the construction of exemplars and networks, particularly in the perception of English nominal suffixes by L1 Thai learners. That is, high-frequency input strengthens the formation of exemplars faster and more effectively (Bybee, 2013). This frequent exposure also facilitates the creation of robust networks, as similar exemplars with shared features are linked within the mental representation. For instance, L1 Thai learners exposed to nominal suffixes like '-ation' and '-or' with high type and token frequency are likely to form stronger exemplar clouds and network connections. Conversely, nominal suffixes with lower frequency, such as '-ence' and '-ure', may result in weaker exemplar formation and less robust networks.

The results from the GJT corroborated with the results from the previous two studies (Chiarakiat, 2019; Sayer & Abdulsalam, 2018), in which input frequency played a positive role in L2 learners' perception of English derivational morphemes. Both studies suggested that frequent exposure promoted mental representation of derivational morphemes, enhancing learners' sensitivity to the target test items with high frequency.

The native group's accurate judgement of all target test items suggested their sensitivity to the target nominal suffixes, which was unaffected by input frequency. This finding aligned with Ellis et al.'s (2008) study, which found that input frequency did not influence native speakers' production of English academic formulaic expressions. The explanation

provided was that native English speakers encountered these expressions so frequently regardless of their degree of frequency that further exposure to the target test items did not significantly improve their ability to differentiate between them. Similarly, in this study, the native English speakers encountered the nominal suffixes so frequently as their L1 that neither high frequency nor low frequency significantly affected their grammaticality judgements.

The Role of Token and Type Frequency on the Grammatical Judgement of English Nominal Suffixes by the L1 Thai Learners in Both Proficiency Groups

It was mentioned in the second hypothesis that token frequency and type frequency played inseparable roles on the acquisition of English nominal suffixes by the L1 Thai learners. Post-hoc analysis using Tukey's HSD was carried out after the one-way ANOVA to examine differences among pairs of means of all four conditions. The results indicated that both type and token frequency significantly influenced the GJT scores, with observed differences between conditions and different proficiency groups and the second hypothesis was partially confirmed.

In the intermediate group, the impact of token frequency was stronger than that of type frequency, as Condition 3 (LH) received the most accurate judgement, followed immediately by Condition 1 (HH), Condition 2 (HL), and Condition 4 (LL). However, no significant differences were found between the conditions. This suggested that the intermediate group was in the developmental phase of English nominal suffixes knowledge and at the intermediate level, the L1 Thai learners might rely more on whole-form storage facilitated by token frequency. Additionally, an affix with high type frequency needs to be encountered to a certain degree for speakers to detect them as an affix (Jarmulowicz, 2002). Speakers must experience several words containing a particular suffix (e.g., 'correspondence', 'dependence', and 'preference' for the nominal suffix '-ence') to increase awareness of productivity. Once productivity increases, the speakers will develop the ability to parse the constructions of these derived nouns. The results from the GJT in the intermediate L1 Thai group were consistent with Chiarakiat (2019), which revealed the dominant role of token frequency in the perception of adjectival suffixes by the intermediate L1 Thai group. This study explained that in the intermediate group, token frequency promoted the underlying representation of words, resulting in faster access as a whole form and a loss of their internal structure.

However, in the advanced group, there was a significant difference between the mean of Condition 1 (HH) and that of Condition 4 (LL) (see

Table 8). That is, the advanced L1 Thai group judged the target test items with high type and token frequency most accurately, and those with low type and low token frequency least accurately. This finding suggested that type and token frequency jointly contributed to nominal suffix acquisition. Token frequency facilitates memory representation (Bybee & Beckner, 2010), while type frequency enhances productivity (MacWhinney, 1978 as cited in Croft & Cruse, 2004). In particular, a linguistic element with high type frequency can be applied to a wide range of new constructions. The results from the advanced group were in accordance with Croft and Cruse's (2004) suggestion that the network structure for token frequency, involving the entrenchment of a linguistic form, and type frequency, involving productivity, collaborate in establishing the mental representation of linguistic forms.

To confirm the hypothesis regarding the combined role of type frequency and token frequency, the mean of Condition 1 (HH), which had high type and token frequency, must be significantly higher than that of Condition 4 (LL), which had low type and token frequency. The results from both proficiency groups partially confirmed the hypothesis. In the intermediate group, the mean of Condition 1 (HH) was not significantly higher than that of Condition 4 (LL) and there were also no significant differences among other pairs. Thus, there was no evidence of collaboration between type and token frequency at this proficiency level. However, the advanced group achieved the highest mean in Condition 1 (HH), which was significantly higher than in Condition 4 (LL). This demonstrated the joint contribution of type frequency and token frequency to the perception of English nominal suffixes by the L1 Thai learners.

Overall, the results provided insights into the developmental trajectory of morphological acquisition. At intermediate levels, the L1 Thai learners were likely to store derived nouns as whole forms, evidenced by their sensitivity to the nominal suffixes with high token frequency. Later, once the L1 Thai learners experienced more novel nouns containing particular nominal suffixes, their parsing skills tend to improve along their proficiency development. This was consistent with the developmental path of the type-token ratio found in longitudinal studies by Yuldashev et al. (2013) and Eskildsen and Cadierno (2007). Based on the Usage-based Theory, it was proposed that L2 learners in early stages are likely to rely on fixed schemas due to the effect of token frequency and later become able to parse and apply specific linguistic items to specific slots in novel constructions due to the effect of type frequency (Eskildsen & Cadierno, 2007; Yuldashev et al., 2013).

The Comparison of the Results from Both Proficiency Groups Together with the Discussion

Results from the Comparison of the Results from Both Proficiency Groups

The independent samples t-test was performed to compare the pairs of the means in each condition from the intermediate and advanced L1 Thai groups. The results are shown in Table 10.

Table 10

Comparison of the Correct Answers on the GJT by the L1 Thai Intermediate and Advanced Groups Across All Four Conditions from the Independent Samples T-test

Condition	Proficiency levels	Number of participants	Mean	*SD*	*df*	*t*	*p*-value
Condition 1 (HH)	Intermediate	30	2.83	.70	58	5.13	<.001***
	Advanced	30	3.63	.49			
Condition 2 (HL)	Intermediate	30	2.50	.18	51.63	4.16	<.001***
	Advanced	30	3.40	.12			
Condition 3 (LH)	Intermediate	30	2.90	.71	58	2.44	.018*
	Advanced	30	3.33	.66			
Condition 4 (LL)	Intermediate	30	2.40	.62	58	3.93	<.001***
	Advanced	30	3.07	.69			

Significance level: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

The independent samples t-test comparing the means of GJT results revealed that the advanced group obtained significantly higher scores than the intermediate group across all four conditions. In Condition 1 (HH), the advanced group's mean (M = 3.63, SD = .49) was significantly higher than that of the intermediate group (M = 2.83, SD = .70), with $t(58) = 5.13$, $p <$ .001. In Condition 2 (HL), the advanced group's mean (M = 3.40, SD = .12) was significantly higher than that of the intermediate group (M = 2.50, SD = .18), with $t(51.63) = 4.16$, $p <$ .001. In Condition 3 (LH), the advanced group's mean (M = 3.33, SD = .66) was significantly higher than that of the intermediate group (M = 2.90, SD = .71), with $t(58) = 2.44$, $p <$ .05. In Condition 4 (LL), the advanced group's mean (M = 3.07, SD = .69) was significantly higher than that of the intermediate group (M = 2.40, SD = .62), with $t(58) = 3.92$, $p <$ .001.

A 2 × 4 two-way mixed ANOVA was utilized to explore the impact of proficiency levels and input frequency on GJT scores. The results are presented in Table 11.

Table 11

Results of the Two-Way Mixed ANOVA Analysis for the GJT

Source	Sum of Square	*df*	Mean Square	*F*	*p-value*	Partial Eta Squared
Proficiency levels	29.400	1	29.400	32.123	.000***	.356
Input frequency	8.483	3	2.828	8.042	.000***	.122
Proficiency levels × Input frequency	1.833	3	0.611	1.738	.161	.029
Between Groups	53.083	58	0.915			
Within Groups	61.183	174	0.352			

Significance level: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

The results revealed a significant main effect for proficiency levels, $F(1, 58) = 32.12$, $p < .001$, $\eta_p^{2\,7} = 0.36$ and a significant main effect for input frequency, $F(3, 174) = 8.04$, $p < .001$, $\eta_p^2 = 0.12$. However, the interaction between proficiency levels and input frequency was non-significant, $F(3, 174) = 1.74$, $p > .05$, $\eta_p^2 = 0.29$. It was revealed that proficiency levels had a greater effect size than that input frequency.

The Discussion on the Relationship Between Input Frequency and Proficiency Levels on the Grammaticality Judgement of English Nominal Suffixes by the L1 Thai Learners in Both Proficiency Groups

The third hypothesis stated that input frequency and proficiency levels interacted in the perception of English nominal suffixes by the L1 Thai learners. The results from the two-way ANOVA analysis revealed that despite the significant main effect of each variable, there was no significant interaction between input frequency and proficiency levels. Each variable independently influenced the learners' grammaticality to English nominal suffixes, suggesting separate mechanisms in the perception of these suffixes. Thus, the third hypothesis was rejected.

The lack of interaction between input frequency and proficiency levels was evident in the consistent patterns of grammaticality judgement across proficiency groups. Both groups showed similar sensitivity to target test items with high input frequency and least sensitivity to those with low input frequency. These findings suggested that input frequency and proficiency levels impacted the perception of English nominal suffixes through distinct pathways. Proficiency refers to language skills (Richards & Schmidt, 2013), while input frequency relates to encountering language patterns (Ellis & Collins, 2009). Proficiency impacts learners' ability to use a language effectively, while input frequency influences the pattern entrenchment in mental storage. The results from this study were in line with Zhao and Le (2016), indicating that the interaction between input frequency

and proficiency levels was non-significant due to their individual effects on the acquisition of English phrasal verbs.

Apart from examining the interaction between input frequency and proficiency levels, it was found that proficiency levels exhibited a greater impact, as evidenced by their larger effect size (see Table 11). The stronger influence of proficiency levels on the perception of English nominal suffixes by L1 Thai learners could be explained by the fact that L2 proficiency levels involve more dimensions of morphological knowledge.

Proficiency levels are directly related to vocabulary knowledge, including form, meaning, and use (Nation, 2001) and associated with L2 learners' ability to recognize both grammatical structures and meanings associated with derivational affixes (Leontjev et al., 2016). Thus, proficiency levels involve more aspects of vocabulary knowledge, affecting L2 learners' perception of English nominal suffixes. The L1 Thai learners with high proficiency are likely to utilize more vocabulary knowledge in judging the accuracy of derived nouns containing English nominal suffixes, resulting in significantly higher means than the intermediate group in all four conditions. In addition to vocabulary knowledge, proficiency levels were found to be associated with derivational awareness (Menut et al., 2023). In other words, as learners become more proficient in their L2, their derivational awareness—the ability to parse derived words into their root forms and affixes—improves (Menut et al., 2023).

Regarding input frequency, it involves the effect of frequent exposure on the entrenchment of the mental schema of a specific form, resulting in stronger recognition (Bybee & Beckner, 2010). Hence, input frequency is mostly related to the entrenchment of the form. Despite encountering specific forms regularly, the L2 learners may not fully develop their vocabulary knowledge if they lack understanding of other aspects such as meaning or function. Thus, more comprehensive vocabulary knowledge, including knowledge of form, meaning, and function was more vital than the retrieval ability of linguistic forms. This resulted in the significantly higher sensitivity of the advanced L1 Thai group to the target test items in all four conditions.

Other Effects

Apart from input frequency and proficiency levels, other factors were observed to influence the perception of English nominal suffixes by the L1 Thai learners.

First, due to the nature of the GJT, which includes both grammatically correct and incorrect items, the L1 Thai learners judged grammatically correct test items more accurately but judged ungrammatical

ones less accurately. Despite recognizing grammaticality, they often struggled to supply the correct nominal suffixes. For instance, in Condition 4 (LL), where 'contingency' was correctly answered by both intermediate and advanced groups, '*apprenticity'—a nonce word for 'apprenticeship'—received lower scores (0.7% for intermediate and 23% for advanced). Interestingly, the advanced group tended to judge ungrammatical items more accurately than their intermediate counterparts, suggesting a reduced task effect with higher proficiency.

Secondly, it was found that the target nominal suffixes with form-similar competitors were judged less accurately. For example, the noun 'chemist', which was grammatically correct in Condition 2 (HL), was judged with 73.33% accuracy by the intermediate group and 90% by the advanced learners. However, some L1 Thai learners judged this word as incorrect and changed it to 'chemistry', resulting in an inaccuracy rate of 26% by the intermediate group and 10% by the advanced learners. This suggested that they were influenced by the similarity between 'chemist' and 'chemistry'. This finding was in line with the neighborhood frequency effect (Grainger et al., 1989), where forms with more frequent similar competitors—or neighbors—are recognized more slowly. Despite 'chemist' being the default grammatically correct test item in this study, some participants judged it incorrectly and answered 'chemistry' instead. This indicated that 'chemistry', with its higher word frequency (198), competed with the target test item 'chemist' (130), even though the context favored 'chemist'. This could be because the participants were more familiar with 'chemistry' due to its higher word frequency.

Third, based on the results of this study, it was evident that regardless of the variation in suffix forms within the same word class (e.g., '-or' vs. '-ist'), the effects of input frequency were apparent. Conditions with higher frequency were judged more accurately than those with lower input frequency. Specifically, the derived nouns containing '-or' ('contractor' and 'moderator') in Condition 1 (HH) were judged accurately at 75% by the intermediate group and 75% by the advanced group. However, the derived nouns containing '-ist' ('chemist' and 'journalist') in Condition 2 (HL), which had high type but low token frequency, were judged less accurately—66.67% by the intermediate group and 86.67% by the advanced group. Thus, despite having similar thematic relations, the nominal suffixes with higher frequency (e.g., '-ation' and '-or') were more likely to be judged accurately compared to the variations with lower frequency.

In addition, the type of roots in each derived word should also be considered. The study found that both proficiency groups judged the target test items containing bound bases (e.g., 'agriculture' and 'chemist') most accurately: 73.33% in the intermediate group and 93.99% in the advanced group. This was followed by the target test items containing adjectival roots

(e.g., 'contingency' and 'efficiency'), which scored 71.67% in the intermediate group and 88.33% in the advanced group. These findings suggested the impact of whole-form representation, influenced by token frequency (Croft & Cruse, 2004). L1 Thai learners were likely to store the whole forms of words like 'agriculture' and 'chemist' rather than their bound bases, which cannot stand alone as words, resulting in the higher accuracy scores. When examining the adjectival roots, which received the second-highest accuracy, it is notable that the words 'contingency' and 'efficiency' were in Condition 4 (LL). This did not necessarily indicate a reduced impact of input frequency. Instead, it might be because their root forms ('contingent' and 'efficient') are orthographically similar to their derived forms, enabling more accurate judgements. This aligned with the findings of Alotaibi and Alotaibi (2017), who noted that neutral morphemes, which do not alter the phonological structure of the stem they are affixed to (Kiparsky, 1982, as cited in Alotaibi & Alotaibi, 2017), are acquired faster than non-neutral suffixes, which result in phonological modifications. In this study, both target test items with adjectival roots contained neutral suffixes, making them easier to acquire. It was found that both proficiency groups scored lower when judging the target test items containing verbal roots: 70.48% for the intermediate group and 84.29% for the advanced group. The lowest scores were observed for the nominal roots, with the intermediate group scoring 56% and the advanced group scoring 73.33%. Despite the high frequency of some derived forms in Condition 1 (HH), those containing verbal roots (e.g., 'alteration' and 'moderator') included non-neutral nominal suffixes such as '-ation' and '-or'. These suffixes possibly posed difficulty for L1 Thai learners in grammaticality judgements. As for nominal suffixes, the derived words were all of low token frequency, as they fell into Condition 2 (HL) or Condition 4 (LL). This limited exposure could contribute to the lower mean scores for grammaticality judgements of the target test items containing nominal roots. However, it should be noted that there were only two target test items containing bound bases and two items with adjectival roots. This increased the likelihood of obtaining higher accuracy scores in these two categories compared to those with nominal roots and verbal roots, which had seven and five derived noun items, respectively.

Overall, the intermediate group scored lower than the advanced group on these tests, indicating greater sensitivity to task effects such as those from the GJT and neighborhood frequency. As proficiency increased, the L1 Thai learners were less likely to be affected by these task-related influences.

Conclusion and Implications

This study was aimed at exploring the role of input frequency, along with proficiency levels, on the perception of English nominal suffixes by L1 Thai learners. Specifically, it examined the grammaticality judgement of intermediate and advanced L1 Thai learners to derived nouns containing nominal suffixes, varying in type and token frequency (high and low) across four conditions of the GJT. The overall results exhibited a significant main effect of both variables, with proficiency levels exerting greater influence. It was also found that at early proficiency stages, token frequency played a stronger role, leading to recognition of the whole forms of the target test items. As proficiency increased, type frequency began to play a role, as learners encountered more forms containing specific suffixes, contributing to the entrenchment of English nominal suffixes in mental representation. This finding corroborated with the type-token ratio discussed by Yuldashev et al. (2013) and Eskildsen and Cadierno (2007).

This study made two main contributions. First, with respect to theoretical implications, it deepened the understanding of how input frequency, including type frequency and token frequency, impacted SLA, particularly in the acquisition of English nominal suffixes. Second, it held pedagogical significance by raising awareness among ELT practitioners and curriculum designers about the role of input frequency in acquiring derivational suffixes. Thus, stakeholders may consider developing appropriate materials, such as incorporating target vocabulary frequently in exampled authentic reading and listening material input, to facilitate the L1 Thai learners' acquisition of English nominal suffixes.

Some limitations should be noted. Firstly, this study utilized a cross-sectional design to compare differences between the intermediate and advanced groups. Hence, further research could benefit from longitudinal studies to explore the developmental trajectory of type and token frequency. Secondly, further studies may include interviews to gain deeper insights from the participants. Third, a timed GJT could be employed in future research to reduce the possible influence from metalinguistic knowledge during task completion.

Acknowledgements

This study was supported by the 90th Anniversary of Chulalongkorn University Scholarship. We would like to express our gratitude to the participants, both L1 Thai and native English speakers, for their time and effort in the data collection process, and to the native experts—Mr. Geoffrey

Berteau, Mr. Teaka Sowaprux, and Mr. Timothy Boundy—for validating the target test items.

About the Authors

Chanikarn Thatchatham: A Ph.D. candidate in English as an International Language Program, Graduate School, Chulalongkorn University, Thailand. Her research interests include second language acquisition, morphology, corpus linguistics, and English language teaching.

Nattama Pongpairoj: An associate professor in English as an International Language Program, Graduate School, Chulalongkorn University, Thailand. Her research interests lie in second language acquisition, particularly L2 representations and cognitive processing.

Endnotes

[1]A stem is the form resulting from the combination of a root morpheme and an affix or some affixes (Fromkin et al., 2018). It is not yet a derived word and allows other affixes to be attached to form a more complex stem. For instance, when the root '-believe' is attached with the suffix '-able', it becomes 'believable'. Later, when a prefix 'un-' is attached to 'believable', it results in the derived word 'unbelievable'. In this case, 'believable' is the stem of 'unbelievable' (Fromkin et al., 2018).

[2] The Chulalongkorn University Test of English Proficiency (CU-TEP) is a standardized test designed to assess the English language skills of candidates for academic and professional purposes. It evaluates main skills including listening, reading comprehension, and writing. The total score ranges from 0 to 120 points, aligning with the Common European Framework of Reference for Languages (CEFR) levels A2 to C1 (Charnchairerk, 2021).

[3] The number in the parentheses represents the frequency count of type frequency.

[4] The number in the parentheses represents the frequency count of token frequency.

[5] The derived words for each type of root were categorized as follows: For **verbal roots**, words derived under Condition 1 (HH) included 'alteration', 'combination', 'contractor', and 'moderator'. Under Condition 3 (LH), the derived words were 'correspondence', 'dependence', and 'closure'. For **nominal roots**, the words derived under Condition 2 (HL) were 'journalist', 'criticism', and 'mechanism'. Under Condition 4 (LL), the derived words were 'partnership' and 'apprenticeship'. For **adjectival roots**,

Condition 4 (LL) included the words contingent and efficient. For **bound bases**, Condition 2 (HL) included 'chemist', while Condition 3 (LH) included 'agriculture'.

[6] A bound base is a bound morpheme that carries the core meaning of a word but cannot stand alone as a word. At the same time, it is neither a prefix nor a suffix (Lieber, 2009), as in 'agri' in 'agriculture'.

[7] Partial Eta Squared (η_p^2) is a measure of effect size used in ANOVA tests to indicate the proportion of the independent variables on the dependent variables (Norouzian & Plonsky, 2018). In this study, a Partial Eta Squared was used to indicate the effect size of the two independent variables (e.g., input frequency and proficiency levels) on the dependent variable, which is the scores of the GJT from the intermediate and advanced L1 Thai groups.

References

Almulla, N. (2015). *The role of frequency in L2 structure accuracy: Acquiring the English infinitive-gerund construction by Arabic L1 speakers* [Master's thesis, California State University, Long Beach]. ProQuest Dissertations Publishing.

Alotaibi, M. A., & Alotaibi, A. M. (2017). On the acquisition of derivational suffixes by Kuwaiti EFL learners. *European Scientific Journal, 13*(17), 223–238. https://doi.org/10.19044/esj.2017.v13n17p223

Bertram, R., Laine, M., & Virkkala, M. (2000). The role of derivational morphology in vocabulary acquisition: Get by with a little help from my morpheme friends. *Scandinavian Journal of Psychology*, *41*(1), 287–296. https://doi.org/10.1111/1467-9450.0020

Bybee, J. L. (2013). Usage-based theory and exemplar representations of constructions. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 49–69). Oxford University Press.

Bybee, J. L., & Beckner, C. (2010). Usage-based theory. In B. Heine & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 827–855). Oxford University Press.

Bybee, J. L., & Thompson, S. (1997). Three frequency effects in syntax. *Berkeley Linguistics Society, 23*(1), 378–388. https://doi.org/10.3765/bls.v23i1.1293

Charnchairerk, C. (2021). Predictive validity of the CU-TEP as an indicator of students' academic achievement in English. *PASAA*, *62*, 29–60. https://doi.org/10.58837/CHULA.PASAA.62.1.2

Chiarakiat, C. (2019). *The role of input frequency on the acquisition of English adjectival suffixes by L1 Thai learners* [Master's thesis, Chulalongkorn University]. Chulalongkorn University Theses and Dissertations (Chula ETD).

Croft, W., & Cruse, D. (2004). *Cognitive linguistics*. Cambridge University Press.

Dąbrowska, E. (2004). *Language, mind and brain*. Edinburgh University Press.

Ellis, N. C., & Collins, L. (2009). Input and second language acquisition: The roles of frequency, form, and function introduction to the special issue. *The Modern Language Journal*, *93*(3), 329–335. https://doi.org/10.1111/j.1540-4781.2009.00893.x

Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, *42*(3), 375–396. https://doi.org/10.1002/j.1545-7249.2008.tb00137.x

Eskildsen, S. W., & Cadierno, T. (2007). Are recurring multi-word expressions really syntactic freezes? Second language acquisition from the perspective of usage-based linguistics. *Collocations and Idioms*, *1*, 86–89.

Fromkin, V., Rodman, R., & Hyams, N. (2018). *Introduction to language* (11th ed.). Cengage Learning.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psychophysics, 45*(3), 189–195. https://doi.org/10.3758/bf03210696

Hamawand, Z. (2011). *Morphology in English: Word formation in cognitive grammar*. Continuum.

Hay, J., & Baayen, H. (2003). Phonotactics, parsing and productivity. *Italian Journal of Linguistics, 15*(1), 99–130.

Iwasaki, S., & Ingkaphirom, P. (2005). *A reference grammar of Thai*. Cambridge University Press.

Jarmulowicz, L. D. (2002). English derivational suffix frequency and children's stress judgments. *Brain and Language*, *81*(1-3), 192–204. https://doi.org/10.1006/brln.2001.2517

Laws, J. C., & Ryder, C. (2014). Getting the measure of derivational morphology in adult speech: A corpus analysis using MorphoQuantics. *Language Studies Working Papers: University of Reading*, *6*, 3–17.

Leontjev, D., Huhta, A., & Mäntylä, K. (2016). Word derivational knowledge and writing proficiency: How do they link? *System, 59*, 73–89. https://doi.org/10.1016/j.system.2016.03.013

Menut, A., Brysbaert, M., & Casalis, S. (2023). Derivational awareness in late bilinguals increases along with proficiency without a clear influence of the suffixes shared with L1. *Bilingualism: Language and Cognition, 26*(1), 138–151. https://doi.org/10.1017/S1366728922000402

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

Norouzian, R., & Plonsky, L. (2017). Eta- and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research*, *34*(2), 257–271. https://doi.org/10.1177/0267658316684904

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hoppe (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137–157). John Benjamins Publishing Company. https://doi.org/10.1075/tsl.45.08pie

Prasithrathsint, A. (1996). Stylistic differentiation of /kaan/ and /khwaam/ nominalization in Standard Thai. In S. Burusphat (Ed.), Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics (Vol. 4, pp. 1206–1216).

Prasithrathsint, A. (1997). The emergence and development of abstract nominalization in Standard Thai. In S. Burusphat & T. Ratitamkul (Eds.), *Southeast Asian Linguistic Studies in Honour of Vichin Panupong* (pp. 179–190).

Prasithrathsint, A. (2005). Nominalization and categorization of verbs in Thai. *Proceedings of SEALS*, 15, 73–81.

Richards, C., & Schmidt, R. (2013). *Longman dictionary of language teaching and applied linguistics* (4 ed.). Routledge.

Sayer, I. M., & Abdulsalam, S. (2018). The acquisition of derivational suffixes by students of English at the University of Human Development. *Journal of University of Human Development*, *4*(4), 55–60. https://doi.org/10.21928/juhd.v4n4y2018.pp55-60

Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *TESOL Quarterly, 36*(2), 145–171. https://doi.org/10.2307/3588328

Smyth, D. (2002). *Thai: An essential Grammar*. Routledge.

Tomasello, M. (2003). *Constructing a Language*. Cambridge University Press.

Wudthayagorn, J. (2018). Mapping the CU-TEP to the Common European Framework of Reference (CEFR). *LEARN Journal: Language Education and Acquisition Research Network, 11*(2), 163–180.

Yuldashev, A., Fernandez, J., & Thorne, S. L. (2013). Second language learners' contiguous and discontiguous multi-word unit use over time. *The Modern Language Journal, 97*(S1), 31–45. https://doi.org/10.1111/j.1540-4781.2012.01420.x

Zhao, H., & Le, F. (2016). Measuring L2 explicit knowledge of English verb-particle constructions: Frequency and semantic transparency at two proficiency levels. In L. Sánchez & R. P. Leow (Eds.), *The Usage-*

Based Study of Language Learning and Multilingualism (1st ed., pp. 171–185). Georgetown University Press.

Zyzik, E. (2009). The role of input revisited: Nativist versus usage-based models. *L2 Journal, 1*(1), 3–17. https://doi.org/10.5070/l2.v1i1.9056

Appendix A

The target test items in the Grammaticality Judgement Test

1. I think the alteration is important to the contract.( ) _____
2. A combinement can be powerful. ( ) _____
3. A contractor is responsible for overseeing the construction of the new building. ( ) _____
4. The moderatist kept the discussion on track during the meeting. ( ) _____
5. She is a chemist and drug safety is her area of expertise. ( ) _____
6. I started my career as a journalator two years ago. ( ) _____
7. When providing criticism, try to focus it on the task, and never the person. ( ) _____
8. Please examine the mechanition carefully before using it. ( ) _____
9. The correspondence included a detailed report on the project. ( ) _____
10. The dependency can lead to difficulties. ( ) _____
11. A closure is often necessary for renovations. ( ) _____
12. The economy of Thailand relies much on agricultory. ( ) _____
13. The partnership resulted in the development of the products. ( ) _____
14. An apprenticity is a great way to gain practical skills. ( ) _____
15. The contingency was planned for in case of an emergency.( ) _____
16. Time management is a key factor in achieving efficience. ( ) _____