Exploring Potential Correlation Between CEFR Grammar Profile in English and Learners’ Overall Grammar Knowledge: An Exploratory Study of “that”

Kasin Janjaroongpak*

College of General Education and Languages, Thai–Nichi Institute of Technology, Thailand

Corresponding Author: Kasin Janjaroongpak, E-mail: kasin@tni.ac.th

ABSTRACT

This preliminary study explored the possibility of using an opaque polysemous grammatical unit as a representation for the whole grammatical knowledge of a learner. There were two groups of informants, a group with a certified language proficiency level at CEFR B1-B2 and a learner group at A1-A2 level. The informants were asked to provide a Thai translation of an English text as faithful as possible. The first group consistently and correctly answered the questions by supplying the right translation of polysemous “that” while the answers from the second group were divided in that some could correctly identify C2 function of the word, “that”, though their overall grammatical knowledge was considered to be at A1 but other A2 students failed to identify C2 function of the word in question. The result indicated that the grammatical construction in question could not be used as a key predictor for learners’ syntactic representation as the lineage relation between CEFR level of the grammatical unit corresponded with the translated texts provided to a limited extent. On pedagogical implications, insights provided suggested that instructors should spend more time explaining challenging advanced grammatical functions as they were points that learners were struggling with and one possible way to check whether they did understand syntactic meaning of a function word was by asking students to supply a translation.

INTRODUCTION

English grammar has been a central piece of investigation for decades, linguistically and pedagogically. A large number of strenuous efforts were put to devise the best possible objective and operationalized explanations both of the grammar itself as well as classroom-oriented implications. Linguistically speaking, English grammar was considered one of the most complex grammatical system in Indo-Aryan language family because of its hybridity, being influenced by French-a Romance language-coupled with its Germanic heritage. One insightful observation (DeKeyser, 2005) contended that there were three pillars of confusion for second-language learners (L2), which were complexity of form, complexity of meaning, and complexity of the form-meaning relationship. These three concepts resided under an overarching term, an opacity. According to (Ellis, Language teaching research and language pedagogy, 2012), form-function was of paramount importance in pedagogical sense in that L2 were struggling with making sense of or mapping between incomprehensible signal, an L2 lexicon, and abstraction it pertained, which might or might not exist in their L1. One case in point would be tense-aspect in English which was not transparent in Thai.

In addition, not only was vagueness confined to syntactic deficiency but also multiple syntactic meaning of a given grammatical unit. According to (Cambridge University, 2005), a function word might have more than 1 meaning. For instance, “by” could be used to express a means of transport in “by bus” while it could convey an amount of something such as “increase by 10 units”. L2 were left to figure out which was which in the midst of a reading passage or listening task. To further complicate the situation, this disparity was divided along proficiency lines.

The struggle was much aggravated when it came to a standardized test. From (Rogers, 2017), an incomplete sentence section in part 5 of TOEIC was testing the ability to discriminate multifunctional nature of function words, identifying the specific function of a particular word. For example, “The person taking the minutes was seated __________ the manager.” and the choices were “A) from B) to C) next D) by.” Obviously, this test specification called for a test taker to demonstrate the extent of his or her detailed knowledge on...
possible functions of several prepositions. This was supported by a landmark piece of research by (Grabe & Stoller, 2011) arguing that grammar was one of determining factors of reading comprehension.

(Croft & Cruse, 2004) (Geeraerts & Cuyckens, 2007) proposed that grammar inevitably involved cognition and translation (Schwieter & Ferreira, 2017) seemed to be a plausible method to tease out mapping mechanism via indirect observation.

In this sense, the aim of this research is to explore the extent could “that” signals correlation between learner’s knowledge over “that” and the CEFR level it pertained?

LITERATURE REVIEW

English grammar, in particular, made use of a limited set of function words or repeated syntactic constructions to express grammatical propositions or meaning. This led to polysemous nature of a plethora of function words (Weibelhuth, 1995) (Noam, 1995) (Radford, 1997). In other words, owing to the constraint that a finite set of these had to perform grammatical expressions as much as possible in one single unit without compromising the accuracy of language use and interpretation. In this vein, some prepositions such as “by” could carry more than three syntactic meanings. However, the meanings pilled on a given functional unit should not be holding two flatly contradictory meanings. One notable example would be “on” in that “on” meaning placing something above a surface or somewhere noticeable should not have a meaning of “off” meaning somewhere hardly being noticed or else it would be too confusing for a child to learn in a short period of time.

To better demonstrate this point, I would like to resort to a corpus-driven dictionary (Cambridge University, 2005). Take determiners as a prime example, the profile went as followed;

The Table 1 illustrated that both determiners and conjunctions had more than 60% of the totals falling under A1 to B1 level, approximately 65% for determiners and 60% for conjunctions. Keeping these function words in perspective, their samples were as followed. For A1 level determiners, among them were “a”, “all”, “any”, “some”, “that”. For the A2 level, there were “another”, “enough”, “few”, “less”. For the B1 level, there were “certain”, “least”, “little”, “other”. For the B2 level, there were “every”, “neither”, “whichever”. For the C1 level, there was “some”.

For conjunctions, A1 level conjunctions were “and”, “because”, “or”. For the A2 level, there were “as”, “before”, “that”, “than”. For the B1 level, there were “after”, “although”, “if”, “yet”. For the B2 level, there were “considering”, “nor”, “whereas”, “while”. For the C1 level, there was “suppose”. For the C2 level, there were “albeit”, “only”. From the lists, there was some overlapping between these two groups, a word such as “that”. This signified that mastering some keywords might be considered particularly advantageous for the learning process.

The fact that a sizeable number of function words fell under A1 to B1 level could be accounted for by observing the benchmark provided by standardized tests. (Tannenbaum & Wylie, 2019) proposed that 550 TOEIC scores were considered B1 level, which was an acceptable level of scores for various internationals in Thailand. This went in accordance with (Cambridge Assessment English, 2018) pointing out that B1 level was an entry level or threshold of operational business English. For IELTS (Cambridge English Language Assessment, 2015), B1 was commensurate with IELTS band scores ranging from 4 to 5, which was an entry level of independent user. Learners at this level were capable of communicating on broad-view level in so doing they, invariably, needed to make use of diverse function words to achieve their communicative purposes.

In particular, the herculean tasks of memorizing and properly producing grammatical units were extremely challenging for L2 as it was perfectly possible for a given function word to carry more than one meaning. To illustrate, “some” had more than one possible meaning. On A1 level, “Some” meant an unknown quantity of something in “There is some cake on the floor.” On B2 level, “Some” meant an amount of something that was of considerable size in “It took some years to master the art.” On C1 level, “some” referred to a specific person or thing without exactly directing to that person or thing in “Some unlucky person was walking down the street.” Obviously, grammatical morpheme or semantics was a deciding factor in determining CEFR level of a given word.

Misinterpreting the specific semantics expression of a given function word could hinder both the overall comprehension as well as linguistics accuracy capability required for completing a test. In this case, should learners be confused with the choices made to decide which meaning of “some” would be assigned to “Some unlucky person was walking down the street.” or they did not know that apart from the typical construction, “some + noun (plural)”, it was possible to have “some + noun (singular)”, expressing different meaning. In the test, a blank test could be used to tease out this feature by leaving the copular verb for test takers to choose. In terms of interlanguage, broadening and developing semantics and syntactic knowledge was considered an integral part of language learning. Students at B1 level were expected to expose to a syntax-sophisticated text in the course of development, meaning that they would read a C1-level text resorting to C1 meaning of “some”. For this, having accurate syntactic representation was particularly vital for learners taking a standardized test.

| Table 1. CEFR level distribution (Determiners and Conjunctions) |
|-----------------|-----------------|-----------------|
| **91 determiners** | **85 conj.** | **57 conj.** |
| A1 | 34 | A1 | 11 |
| A2 | 15 | A2 | 16 |
| B1 | 16 | B1 | 24 |
| B2 | 14 | B2 | 14 |
| C1 | 1 | C1 | 2 |
| C2 | 0 | C2 | 4 |
| Indeterminate | 11 | Indeterminate | 14 |
To closely observe learners’ syntactic assignments required a special method. I would like to resort to translation as a measure, provisional though it seemed, because of practical and linguistic grounds. Linguistically speaking, some syntactic meanings in English was transparent in Thai. On practical front, it was possible to conduct the test under a classroom setting without compromising the research. For instance, “some” on B1 level could be translated as “นวัณณิ” whereas it could be “นวัณณิ” for C1.

Some (Norman, Levihn, & Hedenquist, 2002) may counter the view that grammar should be taught through explicit means and exposure to extensive reading should work a treat but a closer look revealed that nuances and subtlety should be brought into attention in class because function words were carrying little meaning on their own and intermediate learners needed to be scaffolded to start noticing the shades of differences. In contrast, lexis could be left to learners’ autonomous learning, to a certain extent, because learners at this level were capable of searching and verifying their understanding to a certain degree and differences on word meaning could be deciphered from embedded context. As an illustration, “if” could appear in “if” as in reported speech or in conditionals. They both could perform as an opener of a clause such as “I don’t know if you could come” or “If I were you, I would move.” With limited data to analyze, students might not correctly assign the meaning of “if”. At the same time, context helped define the meaning when it came to lexis. This was shown by the following example. “He made a resolution to often visit his aunt.” was startlingly different from “This machine gives us such high resolution that we can see bacteria.” By consulting a good dictionary, learners would know right away that the meaning between these resolutions stood wide apart.

To the best of my knowledge (Hands, 2011) (Brown, 2013) (Klammer, 2013) (Larsen-Freeman & Celce-Murcia, 2016), “That” was considered one of the most versatile grammatical unit in English. Exhaustive list though this table was not, “that” functions could be laid out by the following Table 2.

This function word appeared from A2 to C2 subject to its meaning. Although complete understanding of “that” could help making sense with the content to a large extent because of its ubiquity (Noan, 1995) (Nation, 2016), high-elementary to intermediate learners, A2 to B1, might not fully grasp the meaning of “that” thanks to proficiency level. I speculate that this feature could shed light on the extent of grammatical proficiency level on a small scale.

### METHODOLOGY

An adapted translation test was created according to a Cambridge dictionary (Cambridge University, 2005), the position of “that” were varied, some were attributive while the others were predicative as well as the type of statements, some were affirmative whereas some were interrogative. The item order was randomized to avoid pattern detection.

Purposive sampling was applied in this study to tease out grammatical representation differences between learners of different proficiency level. Participants of this exploratory study were from two groups, the first group was 10 1st year undergraduate engineering students whose TOEIC scores were expected to be below 700, A1-A2, whereas the other was 5 undergraduate engineering students whose TOEIC scores were 700 or more, B1-B2. In this study, the year of undergraduate studies should be excluded as a defining factor because the central focus of this exploratory study is the relation between breath and width of mental grammatical representation and overall linguistic competencies. They were instructed to carefully translate the text containing “that” from English to Thai as accurate as possible. Should they know the precise functions of “that”, they would translate the grammatical morphemes into Thai.

### RESULT

The result showed that from 17 items-one item was deleted from the analysis because of disparity in L1-L2 translation, there were 8 items that were unknown to learners and the rest 10 items were known. For unknown group, C1 were, B1 were 3, and A2 were 3 respectively while for known group C2 were 1, B2 were 3, and A2 were 5. The binary nature, 0 and 1, was used to represent the majority of the informants by resorting to “mode” function in Microsoft Excel.

From quantitative analysis, it noted that “that” was hardly a direct bellwether of a comprehensive overview of learner’s grammar profile. In the other words, one grammatical feature does not necessarily represent the whole picture of learners’ grammatical competencies.

### DISCUSSION

There were several plausible explanations why “that” was not a presentative of the overall understanding, ranging from L1 transfer, interlanguage theory, hierarchy of

<table>
<thead>
<tr>
<th>Function</th>
<th>CEFR level</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determiner</td>
<td>A2</td>
<td>That book is great</td>
</tr>
<tr>
<td>Pronoun</td>
<td>A2</td>
<td>That is what I want</td>
</tr>
<tr>
<td>Relative pronoun</td>
<td>A2</td>
<td>It is a book that many experts recommend.</td>
</tr>
<tr>
<td>Comparative</td>
<td>B1</td>
<td>It is so good that I need to buy 2.</td>
</tr>
<tr>
<td>Complementizer or appositive or adverbial clause</td>
<td>B2</td>
<td>He said that he would not go.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zack holds a belief that love is all around.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nathalie is quite certain that sports are good.</td>
</tr>
<tr>
<td>Intensifier</td>
<td>C2</td>
<td>I don’t think it is that good.</td>
</tr>
</tbody>
</table>
acquisition, anchoring effect, function word opacity, structural complexity for phraseology, idiomatic expressions, frequency, markedness, as well as insights from focus group.

First and foremost, L1 transfer plays a pivotal role in translation or language acquisition as it was a hallmark of cognitive network or mapping in that L2 would perceive and categorize or digest L2 information and later associate each feature with the existing schematic knowledge in L1. One case in point would be the fact that learners could identify the meaning of complementizer with remarkable ease though it was well beyond their overall linguistic competencies owing to the fact that the L2 complementizer had a direct and overt L1 equivalent. Thai adult learners could simply map their Thai to supply the answers, though it was labelled as B2 in CEFR. This compelling justification rang true with the case of determiner.

A large number of learners were developing their interlanguage which made the features present at their level was prominent. Informants showed 5 at A2 level. This showed that they were comfortable with language they had at disposal.

From (Piememann, 2005), the emergence of phrasal expression relating to verbal element came after phrasal expression. This was further supported by the two types that learners answer incorrectly, comparative and intensifier. These two functions of “that” involve other elements in a sentence to form the meaning intended without the other element the meaning would be different as followed;

“He walks so slowly that I cannot wait.”

“*He walks slowly that I cannot wait.”

Other functions which learners could answer were relatively close to each other such as determiners and pronouns as well as complementizer. In particular, although the function of complementizer was labelled as B2 premised on the cognitive complexity between principal clause and subordinate clauses, which learners had to having fully developed embedded clauses (Haegeman & Gueron, 1999) and its operations related to the two immediate constituency rather than the number of units forming the syntactic expression. Learners at A2 level could detect and identify the meaning of the complementizers because they involved only one element while they could not detect the strings of comparative “that”.

In phraseology which was a branch of corpus linguistics (Hunston & Francis, 2000), the longer the string of phrase was the harder to learn. One of the notable examples would be an idiomatic expression, “at the end of the day”, which was perceived as a single unified semantic unit because changing elements in the continuous string might change the meaning like “in the end”. This was a great conundrum in learner’s ability to process each individual unit and stitch it together to realize only one proposition let alone discriminating among expressions appearing relatively the same but the meaning was startlingly different or so similar that learners could not distinguish. For this convincing reason, the number of grammatical units did play a significant role in grammatical-structure acquisition difficulty. From developmental linguistics perspective (Lidz, Snyder, & Pater, 2016), before emergence of phrasal structure, there needs to be a firm ground on unit-based understanding as a building block for broader and more complex constructions or transitional constructions (Ellis, Second language acquisition, 1997) (Ellis, Understanding second language acquisition, 2015) (Supian & Asraf, 2019).

For another thing, the fact that the constructions that learners could not identify were considered an idiomatic expression. It was argued that idioms were perceived as a single unit rather than componential (Biber & Reppen, 2015). In the strictest sense, idioms, especially in English, were criticized as one of the greatest hurdles in learning English because of its individualistic nature in that learners were expected to study and categorize each of them in mental lexicon individually rather than bundling them together. Learners could not figure out the meaning of each element, “so” and “that”, for example, and eventually come up with the idiomatic meaning of the joint construction.

In terms of frequency, the two functions that learners could supply the correct translation were the two most frequent use of “that” in BNC (British National Corpus). Of the total 1,118,985 hits, 318,540 hits of “that” were used as determiners and 658,689 hits were from complementizers, and the rest were of other functions such as intensifiers. Even though “that” as a complementizer was classified as B2 which was well above informants’ proficiency level but as it was ranked the most frequent in BNC learners had had a considerable number of encounters throughout their lives, facilitating associations.

Another point that might explain the error on the meaning of C2 “that” was its markedness. Most of “that” functions were generally associated with noun directly or indirectly as a pronoun or determiner even as a complementizer where it acted as a subordinating conjunction signaling a clause beginning with a noun, most of the time. Therefore, students have formed a close link between “that” and noun, for this reason, they could not assign a proper structural meaning to “that” which was related to an adjective. From the data, it appeared that some informants chose to flout items with intensifier “that” altogether. An unstructured interview (Brinkmann, 2014) suggested that the informants had no clue what to make sense with the function of “that” and they were not comfortable with giving random answers. For answers they were confident enough, they mainly resort to their Thai language to map the meaning between English and Thai before giving the answers, making functions of “that” which were transparent or surfaced on both languages would be translated and vice versa.

As discussed in the literature review, translation could be used as a means to discern grammatical knowledge if the language in question overtly reflects different grammatical functions on forms. This method might be less effective if the translated text did not realize the function on forms. One notable example would be complementizer “that” in Japanese because the Japanese language did not use relative clause in a sense similar to English.

PEDAGOGICAL IMPLICATION

The aim of this study was to explore to what extent CEFR could explain or be used as a predictor to overall grammatical
First of all, according to learning theories (Krashen, 1985) (Berk & Winsler, 1995), instructors were encouraged to explain in detail on a specific function of a function word which was multifunctional or opaque as learners might not decide or verify the right meaning of the function word that well beyond their level especially function words which their meanings were not overt in L1 or those that collocated with a grammatical unit seldom associating with them. One notable practical provisional activity aligned with this concept was to assign a translation task to the class when a large number of students could not answer a comprehension question correctly owing to grammatical concerns on the reading passage. Teachers should explain by compare and contrast between functions generally associated with such function word and the specific functions hindering learners’ comprehension. In the case of “that”, one of the possible ways to illustrate that was to cite examples such as “It isn’t that good.” VS “That bird is nice.” By highlighting the slots, “good” and “bird”, learners would start notice the clear difference which was instrumental in deriving a different syntactic meaning for each. When students came back with the translated text, teachers had another chance to point out the differences of such function word in context as well as consolidate the understanding.

Another point was promoting an extensive exposure to diverse grammatical meanings of function words by means of extensive graded reader as well as periodical monitoring practice sessions. It was a long-established consensus (Shih, 2019) (Grabe & Stoller, 2019) (Faiyaz, 2020) (Sato, 2019) that graded reader helped develop learners’ interlanguage effectively and naturally. If graded readers could label its grammatical features for each book, it would help both teachers who were struggling with finding the right materials to treat certain grammatical problems of learners naturally and learners who sought to familiarize themselves with peculiar grammatical constructions in a creative way. Frequency being a contributing factor in understanding a grammatical unit, learners should maximize their chances of encounters, the more they saw, the stronger the associations developed (Fotos, 2001). The interval grammar practice was not intended to be a practice for the sake of practice but this treatment was meant to be a session to ensure that learners understood the newly-discovered grammatical units by providing pertinent corrective feedback.

REFERENCES

Author Query???
AQ1: Kindly cite Figures 1-3 in text part