

## *The Relativizer That: A Corpus-Based Interlanguage Study*

Supakorn Phoocharoensil, Thammasat University, Thailand

The Asian Conference on Education & International Development 2015  
Official Conference Proceedings

### **Abstract**

The present study explored the use of the relative clause marker *that* in the interlanguage of Thai EFL students, drawing the data from Thai Learner English Corpus (TLEC). As can be seen by the data from a subcorpus of TLEC consisting of intermediate learners' English, different types of grammatical errors on English relative clauses (ERCs) have been discovered. Of all these syntactic deviations, what appears to be the most serious problem for Thai EFL learners relates to the use of resumptive pronouns, regarded as unacceptable in standard written English. Moreover, the findings also accord with the central claim of Keenan and Comrie (1977)'s the Noun Phrase Accessibility Hierarchy in that the easiest ERC types are apparently the subject RC and the direct-object RC respectively, with the least accessible one being the genitive RC. Pedagogically speaking, the subject relative, i.e. the most unmarked type, should be introduced first to students, and the others with further degrees of difficulty instructed later.

Keywords: English relative clause, language corpus, Thai EFL learner, error, The NPAH

**iafor**

The International Academic Forum  
[www.iafor.org](http://www.iafor.org)

## **Introduction**

Research studies in second language learning of English relative clauses (ERCs) have so far shown that the Noun Phrase Accessibility Hierarchy (NPAH) plays a key role in accounting for English learners' development of L2 ERC knowledge (e.g. Amornwongpeeti & Pongpairroj, 2014; Chang, 2010; Chou, 2006; Izumi, 2003; Rattanasak, 2014). However, many studies on how Thai EFL learners acquire ERCs in relation to the NPAH seem to concentrate more on the ERCs introduced by wh-relative markers, e.g. who, whom, which, whose, etc. (Phoocharoensil, 2012, 2014), whereas there exist fewer studies addressing Thai students' acquisition of relative adverbs, e.g. where, when, and why (Phoocharoensil, 2012). Very few studies to date place an emphasis on the relativizer that; many generally investigated that along with wh-relativizers (e.g. Phoocharoensil, 2009).

It is also evident that a good number of studies in EFL contexts elicited data from a variety of instruments, e.g. a sentence combination task (e.g. Doughty, 1991; Izumi, 2003; Phoocharoensil, 2012; Rattanasak, 2014), a sentence interpretation task (e.g. Amornwongpeeti & Pongpairroj, 2014), a grammaticality judgment task (e.g. Amornwongpeeti & Pongpairroj, 2014; Doughty, 1991; Izumi, 2003), a translation task (e.g. Phoocharoensil, 2009; Rattanasak, 2014), and an essay writing task (e.g. Chang, 2004; Phoocharoensil, 2009).

The current study was therefore aimed at an investigation of the relative marker that appearing in Thai EFL learners' interlanguage, with the linguistic data gathered from a learner corpus, i.e. Thai Learner English Corpus (TLEC), which is representative of Thai students' genuine English competence, rather than artificially elicited data given by other aforementioned means.

## **Review of Literature**

### **The Relativizer THAT**

Biber, Johansson, Leech, Conrad, and Finegan (1999) have presented some interesting corpus-informed facts regarding the relativizer that. That is, this relative marker is very prevalent throughout spoken and written registers, e.g. in conversation, news, academic prose, and fiction. That this relative word occurs in a variety of registers is probably ascribed to its ability to occur in different grammatical and meaning contexts. That, to be more precise, is sometimes used in place of who, whom, and which. For example, similar to who, that in (1), in a subject RC position, refers to the human head people, and that in (2), like who and whom, functions as the direct object of the RC, referring to the human head the people. The use of that is preferable in informal spoken English to reduce the formality expressed by whom and possibly to avoid choosing between who and whom (Biber et al., 1999: 615; Carter, McCarthy, Mark & O'Keeffe, 2011).

(1) They're the people that want to buy our house.

(Carter et al., 2011: 451)

(2) They're the people that she met at Jon's party.

(adapted from Carter et al., 2011: 451)

That, in addition, also bears a resemblance to which in that both can occur in the same grammatical function with animate or inanimate non-human heads (Biber et al., 1999). As in (3), that, which refers to the animate head money, is in the RC-subject position, while that in (4) refers to the inanimate non-human head one lesson and appears in the RC-object position.

(3) Don't take money that doesn't belong to you.  
(Carter et al., 2011: 459)

(4) He finally remembers one lesson that his mum had taught him early.  
(Carter et al., 2011: 459)

Apart from occurring in a noun position, that as a relative marker sometimes functions as a relative adverb, i.e., where, when, and why. For instance, that denotes a place and time in (5) and (6) respectively, while it refers to the reason in (7).

(5) That would be the very last place that Marion and I would want to go [to].  
(Biber et al., 1999: 625)

(6) The day that the wall was opened.  
(Biber et al., 1999: 625)

(7) The reason that he was not better known.  
(Biber et al., 1999: 626)

It is worth noticing that the relativizer that is used in a limited fashion since this relative word is prohibited in non-restrictive RCs (NRCs), the preferred choice being a wh-relativizer (Carter et al., 2011; Loock, 2007), as illustrated in (8), where that leads to a grammatically incorrect structure in English, whilst a wh-relativizer like who in (9) is considered appropriate in such a context.

(8) \*The plan, that we discussed yesterday, will be adopted.  
(Cowan, 2008: 438)

(9) The students, who had to take final exams today, are tired.  
(Cowan, 2008: 437)

### **The Noun Phrase Accessibility Hierarchy and ERC Acquisition**

Keenan and Comrie (1977), through investigation into RC accessibility in 50 different languages, demonstrated that there exist six universal RC types; nevertheless, not all types are present in all languages. For example, English, among very few languages that comprise up to six RC types, is richer in RC types than Thai, which has only three types of RCs, namely SU, DO, and IO (Phoocharoensil, 2009). This groundbreaking language universal is known as the Noun Phrase Accessibility Hierarchy (NPAH), and concentrates on the degree of difficulty involved in acquiring each RC type. As indicated in the NPAH, i.e. (10), the subject relative (SU) is claimed to be the least problematic in first language (L1) acquisition due to its lowest degree of markedness. In contrast, the object-of-comparison relative (OCOMP), on the top of

the hierarchy, theoretically is the most difficult for RC acquisition since, as postulated by Keenan and Comrie, this RC type is the most marked.

(10)

- a. That's the man [who ran away]. (SU)
- b. That's the man [whom I saw yesterday]. (DO)
- c. That's the man [to whom I gave the letter]. (IO)
- d. That's the man [whom I was talking about]. (OBL or OPREP)
- e. That's the man [whose sister I know]. (GEN)
- f. That's the man [whom I am taller than]. (OCOMP)

(adapted from Keenan & Comrie, 1977)

Keenan and Comrie (1977) claimed that if there are relative clauses in one language, SU is the most basic RC type. Simply put, it is posited that if a language permits relativization on RC type X, other existing less marked types in the NPAH are anticipated. For example, Thai allows IO relative clauses, and this implies the other two less marked types, i.e., DO and SU (Phoocharoensil, 2009).

Despite the fact that the NPAH was originally meant for L1 RC acquisition, there have been plenty of studies proving that the NPAH is applicable to L2 RC acquisition as well.

Chang (2010), for instance, was partially in line with the NPAH since the RC data from eight graduate students' research proposals exhibited SU as the most common RC type (78%), followed by OPREP (14%), DO (7%), and GEN (1%) in that order. It is also noticeable that OPREP RCs were used with higher frequency than DO RCs, which contradicts the prediction of the NPAH. The study also examined the two types of RCs according to restrictiveness. More specifically, restrictive RCs outnumbered non-restrictive counterparts. Moreover, the learners were explicitly confronted with distinguishing between the two types of RCs. In other words, they showed little awareness of syntactic differences between the two kinds of RCs.

In order to scrutinize Thai EFL learners' acquisition of English non-restrictive RCs (NRCs), Rattanasak (2014) collected data of NRCs using a sentence combination task and a translation task. In accordance with the NPAH, the data from the sentence combination task were consistent with the NPAH in that the learners' NRC acquisition followed the order of difficulty in the NPAH, i.e.  $SU > DO > IO > OPREP > GEN$ . Still, the RCs elicited by the translation task, showed that Thai EFL students' NRC learning was, to some extent, in support of the NPAH prediction, with the exception of GEN, which turned out to be less problematic than IO and OPREP. Rattanasak also reported on the learners' avoidance strategy. In particular, marked NRC types were avoided because the learners were found to impart the same message with more basic types of NRCs.

According to Amornwongpeeti and Pongpairroj (2014), there was a relationship between NRCs and the NPAH. They found that intermediate and advanced undergraduate students speaking L1 Thai, having done a sentence interpretation task and a grammaticality judgment task, by and large acquired NRC types, as well as restrictive RCs, in the order which the NPAH predicts. The NPAH thus accounts for

NRC acquisition as well. As can be seen in the results, NRCs posed more problems for Thai learners than did RRCs.

Essentially, Thai EFL students also produce resumptive pronouns, i.e., a pronoun copy that is redundant with the relative pronoun or relative adverb, such as the subject pronoun *they*, as in \*I've learned that I shouldn't do things that they hurt them., and this is not acceptable in standard English. Phoocharoensil (2011), in line with Phoocharoensil (2009) and Phoocharoensil and Simargool (2010), discovered resumptive pronouns in Thai learners' interlanguage ERCs. According to Phoocharoensil (2011), resumptive pronouns were found in many RC types in high and low proficiency learners' ERCs. To be more specific, low-proficiency students produced more pronoun copies in writing than did their high-proficiency counterparts. It was posited that the higher the level of proficiency, the less the use of pronoun retention since learners having more L2 English exposure are expected to be more aware of ungrammaticality caused by resumptive pronouns, thereby avoiding inappropriately retaining pronouns in ERCs.

According to Braid (1999), resumptive pronouns are viewed as unmarked features across languages, which means there exist more languages in which pronoun retention is a relativization strategy. With respect to markedness and RCs, resumptive pronouns are more likely to occur in interlanguage than marked counterparts, i.e., overt relative markers.

### **Data Collection Technique in the Present Study**

In this study, data were gathered from a subcorpus of Thai Learner English Corpus (TLEC), i.e. the intermediate-learner corpus, compiled by Aroonmanakun, (2009). The corpus data represent written English of university first-year students from two government universities. This subcorpus was chosen because it was supposed to be a clear reflection of authentic data of ERCs used by Thai EFL learners in general, which should lead to some convincing generalizations as to the use of that as a relative marker.

As for ERC selection, the search was aimed at 500 tokens of that from the intermediate-learner corpus. Only the relative marker that was counted, that with other functions, e.g. as a demonstrative, a demonstrative pronoun, or a noun-clause marker, being disregarded. Next, the tokens of that were categorized into two major types according to grammatical functions, i.e. relative pronouns and relative adverbs (Crystal, 2004). Under relative pronouns are the four RC types in the NPAH (Keenan & Comrie, 1977), i.e. SU, DO, OPREP, and GEN. Regarding that as a relative adverb, the focus was on the three main functions, i.e., that denoting a place, a time, and a reason. Furthermore, an error analysis was performed to identify causes of Thai learners' deviant use of ERCs (Ellis, 2008).

### **Major Findings**

#### **The Relative Pronoun That**

Overall, 83.87 % of the ERCs represent that used as a relative pronoun. The ERC data confirmed the NPAH in that what seemed the most accessible or easiest to master for

Thai EFL learners with intermediate proficiency was the SU RC. The more marked RC positions in the NPAH, namely DO, OPREP, and GEN, caused more learning problems. The learners' ERCs comprise different kinds of errors, e.g. the resumptive pronoun, the null-prep construction in the OPREP, and the incorrect relative-word choice in the GEN. The order of difficulty of these ERC types is provided below, where > means 'more accessible or easier than'.

SU > DO > OPREP > GEN

The findings of the order suggested by the present study, which concentrated on only the relativizer that, lent support to Phoocharoensil (2009), exploring ERCs introduced by wh-relativizers, i.e., who, whom, which, whose, and the relative marker that. The main results concerning the order of RC-type difficulty corroborate the NPAH claim in that more marked RC positions are more difficult for RC acquisition, which is consistent with previous studies verifying the NPAH claim (e.g. Amornwongpeeti & Pongpairoj, 2014; Chang, 2010; Izumi, 2003; Pavesi, 1986; Phoocharoensil, 2009; Rattanasak, 2014).

### **The Relative Adverb That**

Only 16.13% of that was employed as a relative adverb. That is, that was used to denote time (11.83 %) the most frequently, followed by place (3.22 %) and reason (1.08%) respectively.

(11) All the time that I have been close to them I think that is the happiest time.

(12) ...around the dormitory to check the place that we had just left whether there were any of my things there.

(13) ...this thing maybe was a reason that she bring me to this school.

That, as indicated in the data, was used in place of when, as in (11), of where, as in (12), and of why, as in (13). It is also worth noting that a Chi-square analysis revealed a statistically significant relationship between Thai EFL intermediate learners' use of that referring to time and their use of that referring to a place,  $\chi^2 = 9.143$ ,  $p = 0.002$ . In a similar vein, there was a statistically significant difference between that referring to a temporal head and that describing the head reason,  $\chi^2 = 16.667$ ,  $p = 0.000$ . However, the relationship between the locative relativizer that and that referring to a reason was shown to be statistically insignificant,  $\chi^2 = 2.000$ ,  $p = 0.157$ .

### **Learners' Errors on ERCs**

Thai EFL learners extensively applied the pronoun-retention strategy. In other words, they produced resumptive pronouns, which are considered ungrammatical in standard English. As can be seen in the data, resumptive pronouns occurred more in the DO RC rather than the SU one, which is less marked. This clearly supports the Resumptive Pronoun Hierarchy (Keenan & Comrie, 1977), which postulates that pronoun retention often arises in marked RC types. Why pronoun retention was absent in the OPREP and the GEN, which are both more marked than the DO, was

probably due to the limited number of these two marked RC types. Examples of resumptive pronouns are given in (14) and (15) below.

(14) Mary is such a good cat that even the dogs around house like her.

(15) ...it is friendship that I keep it in my mind forever.

In (14), the resumptive pronoun *her* is co-referential with the noun phrase *a good cat*, as well as the relative marker *that*. The resumptive pronoun *it* in (15), similarly, is redundant with the relative word *that*, both of which refer to the noun *friendship*.

As regards the reasons why Thai EFL learners use resumptive pronouns, it is interesting to notice that the learners' native tongue does not allow this kind of pronoun copy, nor does L2 English. Consequently, the presence of resumptive pronouns in the Thai learners' interlanguage ERCs should not be attributed to first language transfer. Taking Braidı (1999)'s claim into consideration, we will understand the fact that resumptive pronouns are common in RCs of most languages around the world, and are thus considered an unmarked feature. No matter whether resumptive pronouns are present in learners' L1 or not, they are inclined to employ pronoun retention in L2 RC learning. Put differently, resumptive pronouns exist in interlanguage RCs irrespective of L1 backgrounds, as confirmed by past studies (e.g. Pavesi, 1986; Phoocharoensil, 2009, 2011, 2012 2014; Rattanasak, 2014).

Aside from pronoun retention, the OPREP RC is also what Thai learners are confused over. The key problem concerns the syntactically incomplete structure of RC where a preposition is missing.

(16) but there is just one place that I am very impressed (with) and still want to go back again.

(17) first occupation that I thinks (of) is flight attendant.

The error of the OPREP RC in (16) arises as a consequence of an absence of the preposition *with*. Likewise, the RC in (17) also subcategorizes for the preposition *of* to form a correct preposition stranding construction. These two null-prep structures are examples of interlingual errors, i.e. errors that are caused by learners' mother tongue. Thai EFL learners are perhaps influenced by a lack of OPREP RCs in Thai (Phoocharoensil, 2009). This null-prep construction is very prevalent in intermediate EFL learners' RCs (Sadighi, Parhizgar & Saadat, 2004). As learners' L2 proficiency rises, the occurrences of this interim structure will decline (Phoocharoensil, 2012; Sadighi et.al, 2004). As Odlin (2003) remarked, learners' L2 grammatical constructions, from time to time, exhibit an absence of indispensable prepositions because of native language influence.

## **Conclusion**

The intermediate Thai EFL learners' writing from TLEC has revealed that the relative marker *that* appears with much higher frequency as a relative pronoun (83.87%) than that occurring as a relative adverb (13.13%). The findings, in addition, largely confirm the NPAH as the SU RC was found to be used with the highest degree of

accuracy, and is thus considered the easiest. The DO RC apparently poses fewer problems for Thai learners than the OPREP type. The GEN relative, finally, is viewed as the least accessible. The difficulty order of these RC types, as shown in the corpus data, gives strong support to the principal claim of the NPAH.

The error analysis has shed light on the challenges faced by Thai EFL students in relation to ERC learning. One of the most obvious errors lies in pronoun retention. Resumptive pronouns are often supplied in L2 ERCs, regardless of learners' L1 backgrounds, probably due to their universality and lower level of markedness. Another common deviation found in the corpus is the null-prep construction, where an obligatory preposition is omitted in the OPREP RC.

As it is evident in the research findings that the SU RC is the most accessible type in English, as in many other languages, EFL teachers in Thailand should (re)consider presenting this RC position prior to the other types since students are more likely to acquire this most basic one and become ready for being familiarized with other more marked or advanced RC types. Additionally, due to the fact that resumptive pronouns and null-prep structures prevail in Thai learners' interlanguage ERCs, it is advisable that the prepared ERC lessons, as well as exercises, highlight these problems so as to discourage students from making these grammatically unacceptable constructions.



## References

- Amornwongpeeti, A. & Pongpaiboj, N. (2014). The acquisition of English restrictive and non-restrictive relative clauses by L1 Thai learners. *NIDA Journal of Language and Communication*, 19(21), 1-54.
- Aroonmanakun, W. (2009). *Thai learner English corpus*. Chulalongkorn University. <http://www.arts.chula.ac.th/~ling/TLE/>
- Biber, D., Johansson S., Leech G., Conrad S. & Finegan E. (1999). *Longman grammar of spoken and written English*. Pearson: Singapore.
- Braidi, S. (1999). *The acquisition of second-language syntax*. New York: Arnold.
- Carter, R., McCarthy, M., Mark, G., & O'Keeffe, A. (2011). *English grammar today: An A-Z of spoken and written grammar*. Cambridge: Cambridge University Press.
- Chang, C. (2010). *The use of relative clauses in Taiwanese EFL graduate students' proposal writing*. Master's Thesis. National Yunlin University of Science and Technology.
- Chang, Y. (2004). *Second language relative clause acquisition: An examination of cross-linguistic influences*. National Chung Hsing University.
- Chou, Y. (2006). *Acquisition of English relative clauses by Taiwanese EFL college students*. MA Thesis, National Sun Yet-Sen University.
- Cowan, R. (2008). *The teacher's grammar of English: A course book and reference guide*. Cambridge: Cambridge University Press.
- Crystal, D. (2004). *Making sense of grammar*. Essex: Pearson Education.
- Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition*, 13, 431-469.
- Ellis, R. (2008). *The study of second language acquisition*. Cambridge: Cambridge University Press.
- Izumi, S. (2003). Processing difficulty in comprehension and production of relative clauses by learners of English as a second language. *Language Learning*, 53(2), 285-323.
- Keenan, E. & Comrie, B. (1977). Noun phrase accessibility hierarchy and universal grammar. *Linguistic Inquiry*, 8, 63-99.
- Loock, R. (2007). Appositive relative clauses and their functions in discourse. *Journal of Pragmatics*, 39, 336-362.

Man, L. J. (2007). *An empirical study of Chinese EFL learners' behavior on English relative clauses: Bridging typology and SLA*. Ph.D. dissertation. Shanghai Jiaotong University.

Odlin, T. (2003). Cross-linguistic influence. *Handbook of second language acquisition*. UK: Blackwell Publishing Ltd.

Pavesi, M. (1986). Markedness, discoursal modes, and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition*, 8, 38-55.

Phoocharoensil, S (2009). *A study of English relative clauses in the interlanguage of Thai EFL learners*. Unpublished doctoral dissertation. Chulalongkorn University.

Phoocharoensil, S. (2011). Resumptive pronouns in learner language. *International Journal of Arts and Sciences*, 4(4), 182-193.

Phoocharoensil, S. (2012). On the adverbial relativiser *Where* and prepositional relatives from an interlanguage perspective. *Hong Kong Journal of Applied Linguistics*, 13(2), 85-114.

Phoocharoensil, S. (2014). Errors on the relative marker *WHERE*: Evidence from an EFL learner corpus. *The Southeast Asian Journal of English Language Studies*, 20(1), 1-20.

Phoocharoensil, S. & Simargool, N. (2010). English relativization and learners' problems. *Journal of Pan-Pacific Association of Applied Linguistics*, 14(1), 109-129.

Ramat, A. G. (2000). Typological considerations on second language acquisition. *Studia Linguistica*, 54(2), 123-135.

Rattanasak, S. (2014). *Acquisition of English non-restrictive relative clauses: An interlanguage study of Thai EFL students*. Master's thesis. Thammasat University.

Sadighi, F., Parhizgar, M. R., & Saadat, M. (2004). Preposition pied-piping and preposition stranding constructions in the interlanguage grammar of Iranian EFL learners. *Asian EFL Journal*, 6(4), 1-33.