ABSTRACT
Citation, an indispensible aspect in academic writing, has recently been receiving considerable research attention. This present study empirically delves into citation practices in Master’s theses written by Indonesian student writers. This study employed corpus-based approach. Applying a purposive sampling, there were 7 exemplary theses included in this study. From those theses, it generated 305 citations which then were built as the corpus. A corpus toolkit AntConc3.4.3w was chosen to assist the analysis. Types of citations were analyzed based on Thompson and Tribble’s (2001) framework. Corpus analysis on types of citations revealed that Verb Controlling was more favoured; while Reference form was totally absent. Since citing contributes to successful writers, this study suggests that citation practices particularly types of citations need to be inculcated in the classroom to help the student writers develop citation skill.

Keywords: Citation, Corpus, EFL Theses

A. INTRODUCTION
Citations in academic discourse can be deemed as an indispensible aspect. It implies connections between the writer’s work and the existing scholarship—which is paramount in academic life. From decades ago, analyses of citation, especially focusing on features of citations such as types of citation, have gained considerable attention from previous researchers (Hyland, 2003; Pecorari, 2006; Petrić, 2012; Swales, 1990; Thompson and Tribble, 2001). Employing types of citation, integral/non-integral in this case, is indicative since writers can attribute authors through these forms. Adnan (2004) and Hyland (1999) have argued that the use of integral citation, especially active forms highlights the role of the authors; meanwhile, the use of non-integral citation and passive forms, especially agent-less passive forms, downplays the role of the authors and/or gives greater emphasis on messages.

Especially for novice non-native writers, having a vital role in an academic life, notwithstanding, citations have posed difficulties (e.g. Alwasilah, 2008; Buckingham and Nevile, 1997; Hyland, 2002a; Liu et al., 2016; Loan and Pramoolsook, 2016, 2015; Sun, 2008; Thompson and Ye, 1991). One of the difficulties manifests in novice non-native writers’ tendency to overuse one type of citation (Jalilifar and Dabbi, 2012; Mansourizadeh and Ahmad, 2011). Basthomi (2009: 148), echoing Adnan’s (2004) observation, acknowledges that research activities, citing behavior includes, “have not been mature enough in the Indonesian academic life”. Corroborating Basthomi’s argument, Arsyad (2014) confirms that reviewing previous studies in English have been troublesome for college students and/or new writers in Indonesia. If the difficulties remain, they may hinder the novice non-native writers to perform in a wider academic community, publishing an article for instance (Sun, 2008).

Research focusing on the features of citations such as types of citation which have been employed by novice non-native writers, Master’s students in this case, has
been widely conducted. For example, Jalilifar and Dabbi (2012) conducted investigation to Iranian MA students, Loan and Pramoolsook (2015) to Vietnamese MA students, Manan and Noor (2014) to Malaysian MA students, and Sun (2008) to Chinese MA students. In order to, as White (2004: 112) supports, “put [citation] into a larger framework” to establish theoretical bases, pertaining to Indonesian academic discourse, there has never been any research delving into citation practices despite recent insightful works by Adnan (2004), Mirahayuni (2002), and Safnil (2000). However, the works focused on analyzing citation practiced by Indonesian expert writers in research articles.

Tracing back to what novice non-native writers have practiced, it is Yugianingrum (2008) who investigated citation practices done by Indonesian undergraduate students. With her purposive random sampling, five theses from English literature and five theses from Linguistics, she focused her analysis on thesis structure, citations, and citation behaviors. Supported by secondary data from observation, questionnaire, and interview, she found that the students needed more exposure to cite appropriately. Illuminated by her influential findings, however, the study has only been directed to undergraduate theses.

This study, therefore, extends the framework of citation analysis in periphery, particularly Indonesia, by empirically delving into citation practices focusing on types of citations in the other type of genre, Master’s theses, written by Indonesian student writers. Introduction section some rationale. It, albeit being troublesome for NNES writers (Swales, 1990), is important because it becomes the firstly-read section by readers after abstract section. Moreover, this section is also designed to accommodate information gap which leads to arguments why a study is worth conducting. If this section is not intriguing and convincing, those readers, as a result, will not be motivated to proceed reading the rest.

Types of Citations

Swales (1990: 148-151), one of the pioneers in citation analysis in EAP, in his book explains two kinds of citations, integral and non-integral. Being seen from syntactical features, these two distinctions are easily recognized. The names of cited authors in integral citations appear as the element of sentences; meanwhile in non-integral citations the names are usually between parentheses. Table 1, retyped verbatim from Swales (1990: 149), epitomizes detail information.

As seen from Table 1, the integral citations show the name of the writer as subject (Ia), passive agent (Ib), as part of possessive noun phrase (Ic and d) and as an adjunct of reporting (Ie). The non-integral citations show three parenthetical citations and two superscripted ones. In Indonesian academic discourse, most of non-integral citations appear in parenthetical forms. To distinguish integral and non-integral citation, according to Safnil (2000: 141), we can identify through the appearance of ‘complementizer that’ (see also Charles, 2006). The +R citations are reporting which means that the writers employ reporting verbs (show, establish, claim) to introduce previous researchers and their findings. Below it, -R citations are non-reporting which means that the writers do not employ reporting verbs.
Table 1 Types of Citation from Swales (1990: 149)

<table>
<thead>
<tr>
<th>Integral</th>
<th>Non-integral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ia</td>
<td>Na</td>
</tr>
<tr>
<td>Brie (1988) showed that the moon is made of cheese.</td>
<td>Previous research has shown that the moon is made of cheese (Brie, 1988).</td>
</tr>
<tr>
<td>Ib</td>
<td>Nb</td>
</tr>
<tr>
<td>The moon’s cheesy composition was established by Brie (1988).</td>
<td>It has been shown that the moon is made of cheese (Brie, 1988).</td>
</tr>
<tr>
<td>Ic</td>
<td>Nc</td>
</tr>
<tr>
<td>Brie’s theory (1988) claims that the moon is made of cheese.</td>
<td>It has been established that the moon is made of cheese.</td>
</tr>
<tr>
<td>Id</td>
<td>Nd</td>
</tr>
<tr>
<td>Brie’s (1988) theory of lunar composition has general support.</td>
<td>The moon is probably made of cheese (Brie, 1988).</td>
</tr>
<tr>
<td>le</td>
<td>Ne</td>
</tr>
<tr>
<td>According to Brie (1988), the moon is made of cheese.</td>
<td>The moon may be made of cheese.</td>
</tr>
</tbody>
</table>

The four distinctions made by Swales above have been extended by Thompson and Tribble (2001). They analyzed dissertations from two departments, Agricultural Botany and Agricultural Economics. From the analysis, they develop a set of categories which combines linguistic (syntactical) aspects and content criteria as drawn in Table 2. The combination is crucial in identifying rhetorical/communicative purposes of citations.

Table 2 Types of Citation from Thompson and Tribble (2001)

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Integral citations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb controlling</td>
<td>The citation acts as the agent that controls a verb, in active or passive voice.</td>
<td>David and Olson (1985) define a management information system more precisely as…</td>
</tr>
<tr>
<td>Naming</td>
<td>The citation is a noun phrase or a part of a noun phrase.</td>
<td>Typical price elasticities of demand for poultry products in Canada, Germany and the UK are shown in Harling and Thompson (1983).</td>
</tr>
<tr>
<td>Non-citation</td>
<td>A reference to another author but the name is given without year reference.</td>
<td>The “classical” form of the disease, described by Marek, causes significant mortality.</td>
</tr>
</tbody>
</table>

**Non-integral citations**

<table>
<thead>
<tr>
<th>Source</th>
<th>Citation functions to attribute a proposition to another author.</th>
<th>Citation is central … because it can provide</th>
</tr>
</thead>
</table>
Identification

It identifies an agent within the sentence it refers to. A simulation model has therefore been developed to incorporate all the important features in the population dynamics (Potts, 1980).

Reference

It is usually signaled by the inclusion of the directive “see”. DFID has changed its policy recently with regard to ELT (see DFID, 1998).

Origin

It indicates the originator of a concept or a product. The software package used was Wordsmith Tools (Scott, 1996).

B. METHOD

This study employed corpus linguistics as the research method with purposive sampling. A particular corpus toolkit chosen to assist data analysis was AntConc3.4.3w. Since the data sources were in written forms, they needed to be converted into .txt file. To convert the forms, ‘OCR to Word’, which is freely downloadable, was chosen.

The corpus was built from seven introduction chapters taken from seven exemplary Master’s theses. The theses were chosen under the following criteria: they are (1) the theses are those in the field for which the writer’s skill is adequate to ensure the analysis; (2) the theses are considered as exemplary ones to establish transferability; (3) the theses are written in English by NNES Master’s students which have been approved by the board of examiners to satisfy the focus of this study, that is citation written in English by NNES Master’s students. From those chapters, there were 305 citations.

In accords with the quality of theses, one social construct of exemplary theses can be based on high-rated theses, that is, those which gained grade A. The data were then analyzed by using Thompson and Tribble’ (2001) framework.

The seven introduction chapters in soft files then were analyzed through the following steps

Step 1 **Identifying the types of citations**
The writer inserted two digits such as 19 and 20 in AntConc3.4.3w; and this software can sort out years of publication written in the corpus. ‘19’ refers to information within the twentieth century (from 1901-1999); meanwhile ‘20’ refers to information within the twenty first century (from 2000). From this KWIC, canonical citations were collected, that is, those with brackets, authors’ names, year, and pages.

Step 2 **Categorizing the types of citations and counting the frequency**
The results from the concordance were then copied in the following Table, which is adapted from Thompson and Tribble (2001), and were
manually categorized based on the framework. Done with categorizing, the writer counted the frequency of each type of citations then calculate the total numbers.

Step 3  
**Summarizing the frequency and calculating the percentage**  
After all of the theses were categorized, the writer sum up the frequency in the following table and calculate the percentage of each type based on the following formula.

### C. FINDINGS AND DISCUSSION OF RESEARCH RESULTS

The corpus was analyzed based on types of citations developed by Thompson and Tribble (2001). Detail information about the result of the analysis is presented below.

#### Table 3 Types of Citations in the Corpus

<table>
<thead>
<tr>
<th>Theses</th>
<th>Types of citation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Integral</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VC</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>37</td>
<td>5</td>
</tr>
<tr>
<td>T2</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>T3</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>T4</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>T5</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>T6</td>
<td>34</td>
<td>5</td>
</tr>
<tr>
<td>T7</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>147</td>
<td>25</td>
</tr>
<tr>
<td>%</td>
<td>48.2</td>
<td>8.2</td>
</tr>
</tbody>
</table>

Table 3 above summarizes quantitative analysis of types of citations found in the corpus. As can be seen from the data in Table 3, a majority of the student writers in the corpus appear to favour Integral citations nearly twice as many Non-integral citations. To be more detail, Verb Controlling is favoured twenty percent more than Source form. However, it needs to be noted that Reference form is very likely less favoured compared to the other forms.

Within Integral citation type, Verb Controlling gained the highest frequency (48.2%). The choice of this syntactical form indicates that the student writers in the corpus gave greater prominence to the authors whose work they cited. As the highest favored, Verb Controlling was found in the form of either active (Excerpt 1) or passive sentences (Excerpt 2). However, active forms far surpass passive occurrences.

**Excerpt 1**

Harmer (2007 : 146) adds that in deciding which textbook to use, teacher can analyze its content appropriateness to the students, the right kinds of activity, and the logical sequences. (T6, underline added)

**Excerpt 2**

It is proved by Umiyatun (2011) in her research that grammar, content, organization, vocabulary, punctuation, and spelling (mechanics) belong to the elements in writing in the
target language. (T2, underline added)

The pattern of the active sentences are either S+V or S+adv+V. Although the latter occurred only once, the role of adverb here, if the student writers strategically employed, can be very useful for giving evaluation to cited works (see Buckingham and Neville, 1997).

Excerpt 3
Krashen (2003: 89) simply states (that) by reading, the students are building vocabulary, improving writing style, and developing grammatical competence (T3, underline added)

Verb Controlling form mostly contains agent by mentioning the names of the authors (proper nouns). The agent can be from primary sources (Excerpt 4) or secondary ones (Excerpt 5). The inclusion of secondary sources may reflect that the student writers in the corpus had limited access to some primary sources to cite at the time they wrote theses.

Excerpt 4
Further, Gonzales (2011: 8) found that the students with the higher score in her classes; enjoyed and used this strategy in their reading comprehension lessons. (T3, underline added)

Excerpt 5
In line with the statements, Widdowson in Mukminatien (1991) states that switching the focus of attention in teaching English from the grammatical to communicative properties of the language is important. (T2, underline added)

Following sentences with proper nouns, the student writers often included subject pronouns referring to those names (pronominal reference). This can be included in non-citation. The sentences with subject pronouns contain extended explanation of the previous Verb Controlling forms.

Excerpt 6
This is supported by Soozandehfar and Noroozisiam (2011) who did similar study on the relationship between cognitive styles and L2 speaking performance. They found that there is no significant correlation between cognitive styles and students’ L2 speaking performance. (T5, underline added)

Although Verb Controlling was the most frequently employed in total, there were two student writers, that is T4 and T5, who less favoured the type. They appeared to prefer Source than Verb Controlling. This unexpected result might be caused by the writing style of the student writers. However, the relationship between the choice of types of citations and writing style, which can also be directed to student writers’ writing competence, do not become the focus of this study. Therefore, future researchers may see the relationship to enrich the discussion of various choices of types of citations.
Naming category was the second most frequent Integral citation used by the student writers in the corpus. In this case, the writer intentionally differed between naming (8.2%) and adjunct reporting (2.3%). The latter will be reported afterwards. According to Thompson and Tribble (2001), there are three sub-categories under naming category: (1) noun phrases which signify text rather than human agent, (2) noun phrases which identify a particular equation, method, formulation or similar construct with individual researchers, (3) noun phrases which refer generally to the work or findings of particular researchers. From 25 (8.2%) Naming form found in the corpus, 9 were in the form of example 1 (Excerpt 7), 3 were in the form of example 2 (Excerpt 8), and 13 were in the form of example 3 (Excerpt 9). It means that the student writers tended to use Naming form in order to present the findings from previous studies.

**Excerpt 7**

Horwitz’s (2001, cited in Baralt & Gurzynski-Weiss, 2011:201) statement confirms that the construct of anxiety is often believed to be the affective factor with the greatest potential to pervasively affect the learning process. (T1, underline added)

**Excerpt 8**

The name of weekly journal is made by the researcher based on a writing technique stated by Raimes (1983). (T2, underline added)

**Excerpt 9**

In the study conducted by Murphy, Casey, Day, & Young (1997), it was figured out that FI learners are better than FD ones. (T5, underline added)

Under adjunct reporting, the student writers tended to only used one pattern, that is ‘according to’ compared to Iranian master students (Jalilifar and Dabbi, 2012).

**Excerpt 10**

One type of favorite feedback given to the students' writing according to Leki (1990, as cited by Hyland (2003)) is feedback on grammar in which the teacher gives marked on the students' errors. (T2, underline added)

Non-citation category was the lowest Integral citation favored by the student writers in the corpus. It took only 3% occurrences. From the corpora, the tendency of using non-citation category was in 4 forms. They are: (1) mentioning a possessive noun phrase which has been known as factive knowledge (Excerpt 11), (2) mentioning proper noun which has been mentioned before, either one single agent or a compilation of some agents (Excerpt 12), (3) mentioning pronoun which refers to previously mentioned agents (Excerpt 13), (4) mentioning a proper noun which has
not been mentioned before but it seems like everyone has a background knowledge about the person (Excerpt 14).

Excerpt 11
Sellers (2000)
strengthens the general drift of Krashen's belief that the more highly anxious students tended to recall less passage content than did those participants who claimed to experience minimal anxiety. (T1, underline added)

Excerpt 12
As reading anxiety and self-efficacy, self-efficacy and reading strategy, and self-efficacy and reading comprehension are reported by the previous researchers have a significant correlation. (T1, underline added) (compilation of some agents)

Cognitive styles are information processing habits representing the learners' typical model of perceiving, thinking, problem solving, and remembering (Keefe, 1987:7; Saville-Troike, 2006:87). Keefe specifically states that "cognitive styles are related to, but different from, intellectual abilities". (T5, underline added) (one agent)

Excerpt 13
In spite of this fact, so many researchers have conducted studies about reading anxiety, self-efficacy, and use of reading strategies in each. However, only few of them have explored the relationship among reading anxiety, self-efficacy, use of reading strategies, and reading comprehension of EFL learners. (T1, underline added)

Excerpt 14
As the culmination, Indonesian Minister of Education and Culture, Anies Baswedan, releases an official missive no. 179342/MPK/KR/2014 about the implementation of Curriculum 2013 in December 5, 2014 to bend its implementation, (T4, undeline added)

Non-integral citation was less favored by the student writers in the corpus. Under some sub-categories, precisely, the result showed that Source citation gained the highest usage; meanwhile, Reference citation in the corpora was absent.

Source citation held around 29.5% occurrence. Mostly, citation in this form contained factive knowledge which reflects that the student writers were in the agreement. The use of source citation in the corpus is mostly to provide background knowledge for their readers in order to establish a territory.

Excerpt 15
The reading difficulties come from different sources: poor interpretation of the texts, poor vocabulary, the use of inappropriate reading strategies, and poor grammatical competence (Behroozizad & Bakhtiyarzadeh, 2012 cited in Hamra & Syatriana, 2012:1). (T1)
A journal is a personal record in the form of written
text in notebook of the everyday events of someone's life and issues that concern
him/her, not intended for public display (Thomkins, 1991; Stanley et al., 1988). (T2)

Identification citations occurred one third less than Source citations. The
citations under this sub-category mostly contained reporting verbs in either active or
passive form. This reflected a basic question “who do that?” which reflects that there
is an agent who do the activity. The student writers seem to give more emphasis to
the message rather than to the author’s attribution.

Excerpt 16
Researchers in the field second of and foreign language teaching correlate
individual differences in second or foreign language learning into three categories:
cognitive factors, affective factors, and miscellaneous factors (Al-Shboul et al.,
2013). (T1, undetline added) (who correlate?)

A number of collaborative classroom action research give the empirical evidences
that SQ3R can improve students' achievement in reading comprehension
(Indahyanti, 2008; Setyawan, 2010; Prasetyo, 2011; Khasanah, 2011; Pribadi,
2013). (T5, underline added) (whose research give the empirical evidences?)

Regarding the practice of
teaching L2 content-area reading, it was revealed that the traditional teaching of
reading in the form of to review and translate the vocabulary, read aloud the text,
and answer comprehension questions did not help the students improve their
reading ability (Farrell, 2001). (T5, underline added) (who revealed that?)

Anxiety is claimed to have
potential negative effects on academic achievement, social context, and
personality of the language learner (MacIntyre, 1999; Kuru-Gonen, 2009;
Jafarigohar & Beehroznia, 2012). (T1, underline added) (who claims?)

Discussion
The findings revealed that Integral-Verb Controlling form was highly favoured
by the student writers in the corpus. The result supports some studies on citation
practices in periphery setting such as Jalilifar and Dabbi (2012), Loan and
Pramoolsook (2015), Mansourizadeh and Ahmad (2011), and Sun (2008) which
confirm that non-native student writers have a tendency to employ Verb Controlling
to cite information in Introduction section in their Master’s theses. The choice remains
preferable since Verb Controlling becomes the easiest and the most obvious form to
practice for non-native student writers (Sun, 2008). Jalilifar and Dabbi (2012) argue
that the preference may be rooted from little awareness of citation conventions in the
center setting, which is also supported by Loan and Pramoolsook (2015). If the
tendency runs on or even overuses, as in Mansourizadeh and Ahmad’s (2011) study,
it, later on, may lead to possible problems such as losing readers’ interest (Sun, 2008)
or gaining little space for publication (Jalilifar and Dabbi, 2012).
The finding is also similar to the practices in the center setting. Charles (2006), for instance, found that native student writers frequently employed Integral Verb Controlling form in their theses. This finding may lead to a temporary notion that in terms of types of citations, native and non-native student writers are likely to perform a similar act—they prefer Verb Controlling form than the other forms. The first reason why they act similar way in terms of employing types of citations despite their being different background is for attribution. Hyland (1999) points out that by mentioning human agents, name of authors in this case, the citing writers give more prominence to the authors than the messages. In relation to advancing their arguments, this step is considered important for student writers (see third reason).

The second reason may be because of the typicality of soft disciplines. Corpus in this study and Charles’ together was taken from soft disciplines, Education and Politics/international relation respectively. According to Hyland (1999), the citing writers within soft sciences tends to employ Integral citations while those from hard sciences do the other. Since knowledge in soft disciplines can be recursively revisited, there is a need to construct knowledge to direct readers to the intended meaning, that is the importance of the research topic. Verb Controlling can be deemed as a strategic tool to construct knowledge since the form “signals human agency in knowledge construction” (Hu and Wang, 2013: 23). Moreover, authors are presented as actively communicating through the use of reporting verbs (Buckingham and Nevile, 1997).

The third reason can be because the student writers want to align with a particular discourse community. The alignment then is useful in advancing their arguments. The notion can be explained as the following. Referring to Swales (1990: 24-27), a group of individuals can be identified as a discourse community if it has six characteristics: there are common goals, participatory mechanisms, information exchange, community specific genres, a highly specialized terminology and a level of members. If the writer departs from the notion of ‘a level of members’ within discourse community, she can say that, members, be they novices or experts, in a discourse community are those who have the same interest in a particular topic and have mechanisms to communicate in order to exchange information through specific lexis and terminologies. Regarding the level of members, student writers can be considered as novices. As the nature of newbie, we—to include the writer herself—have a penchant to ‘passively follow’ instead of ‘actively exchange’ information which we are interested in (later, the writer will explain that student writers neutrally agree the followed information). With this background in mind, mentioning the names of the authors can help student writers engage members of the community whose thoughts—either concepts, explanations or definition of key terms—they follow. The practice then is used to give their statement greater authority in order to backup their claims. According to Sun (2008: 14), giving prominence to the author can reflect that student writers “depend large degree on the authority represented by the authors” to help them boost their arguments.

Another important point to discuss is the use of Non-integral Source and Reference form. Within Source category, the student writers in the corpus tended to include not only one author but also more than that. This reflects that the student writers in the corpus had already had the ability to elaborate some references which
have similar ideas. This is also in line with the knowledge development of graduate students. However, the ability was not completed by the skill to strategically employ, borrowing Petric’s (2007) classification, exemplification by adding “i.e. or e.g”. If it happened, the readers might assume that the student writers had quite huge range of reading materials related to the topic brought on. This is also a good start to introduce the research topics. As an alternative, the student writers could also strategically employ Reference citations, which were absent in the corpus, by adding “see” to direct the readers to refer to related sources which are too lengthy to include. The absence of Reference citations may occur because theses, compared to articles, are considered longer. As a result, theses writers can include longer explanation. In addition, according to Thompson and Tribble (2001), Reference form usually appears within Method section—this study focused on Introduction section. Another reason can be because of their real readers who are their advisors and examiners. Having readers who are ‘higher’ than them, student writers presumably feel impolite to apply Reference form which includes directives such as ‘see’ (Hyland, 2002b). However, this sub-category, if strategically employed, can reflect that the student writers are quite good in separating which citations should be attached within their texts and which ones should be excluded (Petrić, 2007).

D. CONCLUSION
This present study has investigated how Indonesian Master’s student writers practiced citations in the EFL theses. Regarding types of citations, Verb Controlling was highly favored; meanwhile Reference was totally absent in the corpora. The writer has suggested that the preference is presumably affected by the nature of theses, the citation acquisition, the typicality of soft discipline, the position of the student writers to their readers, and the status of being an apprentice in a particular discourse community. Pedagogical implications of this study can be directed to inculcating citation skill for non-native student writers. By explicitly teach how to cite, introducing types of citations for instance, non-native student writers, at least, become aware of citation conventions, which is useful for their academic career.

E. REFERENCES


