1. Introduction
With the continuous development of English language teaching (ELT), language awareness has been studied as a teaching approach. Language awareness means the learner’s development of intensifying their consciousness and sensitivity to language forms and functions (Carter, 2003). This article offers a lexical variation view on language awareness, and attempts to discuss the variability and flexibility with a focus on lexis, and reflects lexical awareness on language teaching from the perspective of a learner, a user, a teacher. The main body parts can be divided into three sections. The first section begins by the definition of language awareness and the broad aspect of language use and knowledge that influenced by different factors; the second section emphasizes on the lexical awareness and discusses the variability and flexibility affected by medium, user, and effectiveness, and how lexical awareness implicit on Data-driven learning (DDL) to EFL learners; section three demonstrates reflection from the perspective of English language teachers.

2. Language Awareness
For one thing, language is viewed as an autonomous system; for the other, it is inseparable from social and contextual factors (Graddol et al., 1994). Generally, the widely used definition of language awareness (LA) is put forward by Arndt, Harvey, and Nuttall (2000): “Language awareness is an approach to language study which draws upon a number of disciplines including language teaching, applied linguistics, and other related areas such as psycholinguistics, sociolinguistics and anthropological linguistics” (p. 12), which is mainly about three types of contexts: language learning in school and higher education, teaching and learning second or foreign languages, and the study of socio-cultural influences upon people’s use of language. They believe that unified principle behind all these multiple approaches of language study is to raise awareness as the goal by examining the characteristics of language and studying how language works (ibid.). Except for these types of contexts what Arndt, Harvey, and Nuttall (2000) mention above, Carter (2003) even considers mother-tongue education, or is not merely including second or foreign languages education. Some authors argue that the term ‘language knowledge’ (Carter, 2003, p. 64) and ‘language about language’ have sometimes been more popular than ‘language awareness’ (Fairclough, 2014, p.1). Hawkins (1984) states that the term ‘language awareness’ is used in the field of education since the early 1980s, which is the new language awareness element that is proposed by language educators in the school curriculum. Fairclough (2014) also thinks that language awareness is concerned about education, making language knowledge an important part of language education. Furthermore, the definition of language awareness by Bolitho and Tomlinson (1995, p. iv) as cited in Bolitho et al. (2003) report that language awareness is helping to cultivate ‘a healthy spirit of enquiry’ and building the classroom into a place where ‘the only views of language that matter are the ones that teachers and learners have built up in their heads’. But, Bolitho et al. (2003) and Tomlinson (1994) admit that language awareness is explored by learners themselves. It is not taught by teachers or textbooks but developed by learners, and it is the internal, progressive, and realization of the use of language in reality, which is driven by curious learners consciously focusing on real instances of language, trying to discover and express patterns of language use, learners ‘discover language for themselves’ (p. 251), and as a way of ‘helping learners to help themselves’ (p. 257) (Bolitho et al., 2003). And Tomlinson (1994) suggests that language awareness is ‘gradually developed internally by the learner’ (p.123).
While other researchers provide a definition of language awareness more squint towards psychology and motivation, not just focus on the aspect of education. In an analysis of ten questions about language awareness (Bolitho et al.) in 2003, Tomlinson defines language awareness as “a mental attribute which develops through paying motivated attention to language in use, and which enables language learners to gradually gain insights into how languages work.” It is also a pedagogic approach that aims to help learners to gain such insights” (p. 251). It is related to psychology and mentality. Likewise, Little (1997) summarises the meaning of language awareness from Nicholas (1991) that it has been used in psycholinguistic to indicate that language awareness is learners’ awareness of language instead of conscious reflection on language. The author (Little, 1997) convinces that the language awareness in the sense of psycholinguistics is part of our innate ability to acquire and deal with language, and the more general educational sense of language awareness is the knowledge of language imparted to learners through school education.

Consequently, language awareness is closely related to language teaching and learning and is with regard to ESL/EFL (English as a second language/English as a foreign language) context. Whomever language teachers and language learners could benefit from language awareness. For instance, for language teachers, language awareness offers teachers to think about what should be included when they try to deliver L1 skills or strategies to other languages (Arndt, Harvey, and Nuttall, 2000); for language learners, it helps learners to form good language awareness and speak in a right way in different situation.

Through various definitions of language awareness from different scholars, what arouses my attention most is that “person’s conscious attention to language or culture and, importantly, their engagement with these” (Byram, 2012, p. 6). This is because language awareness not merely involves language teachers and learners, but concerns everyone in daily life. Everyone has to articulate or write more or less in their lives, and different people may express or write in different ways or types of language due to their discrepant living and working environment, contexts, and educational backgrounds. For example, college students are required to write assignments or dissertation in academic styles, such as the sentence pattern: the studies have reported, while some of the trendy teenagers are keen to use network parlance (internet slang), such as ‘LOL’ expressing laugh out loud (“Internet slang”, 2020). People may be careless about this, but it does take place in everyone and they are participants in this activity. In this paper, language awareness could be understood as people are consciously concerned about the use of language in every aspect of their lives, whatever in written and in spoken. People ought to have a certain degree of conscious or unconscious response to the properties of language that they use.

According to Hales (1997), language awareness could be interpreted as sensitivity to grammatical, lexical, or phonology features, and the effect of different forms of use on meaning. From here, language awareness could be semantically divided into four aspects: phonological awareness, lexical awareness, grammatical awareness, and possibly discourse awareness, whatever in spoken and written except for phonology (only concerning with spoken). For language teachers, the major elements of phonological awareness concluded by Treiman & Zukowski (2013) according to Morais, Alegria, & Content (1987): “awareness of phonological strings (a global, nonanalytical level of awareness); awareness of syllables; awareness of phonemes (also called segmental awareness); and awareness of phonetic features” (p. 6). Phonology is as grammar of phonetic patterns. For example, in English, teachers ought to know that the sounds [b], [d], [e] could only form in these order [bed], [deb], and [ebd] (ebbed); on the contrary, [bde], [deb] and [ebd] are impossible, because [bd] or [db] can not be begun with words, or end words with [db], although they are pronounceable (Atitchison, 1994). The rules of phonology vary by different languages. Yet, the sequence [bd] was permitted at the beginning sound of words, as in bdelurous (means rascal in English) in ancient Greek (Atitchison, 1994). Aşık et al. (2015) claims that “According to Common European Framework of Reference (2001), lexical competence consists of lexical and grammatical elements. Lexical elements include fixed expressions (sentential formulae, phrasal idioms, fixed frames, other fixed phrases such as compound prepositions, fixed collocations) and single word forms (members of the open word classes and closed lexical sets such as days of the week, etc.). Closed word classes (articles, quantifiers, demonstratives, personal pronouns, etc.) belong to grammatical elements” (p. 87). Lükica (2011) expresses that discourse structure awareness is seen as a meta-linguistic awareness, which also includes the knowledge of reading strategies and their appropriate use to address reading problems and correctly interpret text information.

Arndt et al. (2000) and Tomlinson (1994) claim that language or language awareness is dynamic. “It reflects and is reflected by the changing ways human societies order themselves; it influences and is influenced by people’s relationships, activities, and communications” (Arndt et al., 2000, p. 18). They maintain that the effectiveness of language lies in ‘a whole range of extra-linguistic factors’, such as the response, interpretation, and attitude of the receiver of the message (ibid.). A number of elements affect the use of language and language. Arndt et al. (2000) make a summary of a basic framework for awareness of language-in-use with a diagram in the book Alive to Language, and those aspects in the framework are categorized as four layers and crosscorrelated with each other: the central part of diagram is language-in-use that could contain sounds, words, and structures (texts); then is flexibility, and choice of words and structures, which the former supports the latter, because language is dynamic, and enable to adjust various circumstances; the third layer consists of four features: attitudes, effectiveness, medium, and variety; the fourth layer are context of language use and knowledge of the world. These features could affect people’s choice and use of language to varying degrees and could be explained one by one. With the development of technology, new words occur with new meanings on the screen (computer, television, etc.), such as internet slang what discuss above, which shows the flexibility of language. Structure is concerning language organization and structure, and reveals ‘how language work and how the various components interact with each other’ (Arndt et al., 2000, p. 20). Grammar, vocabulary, and pronunciation are all essential elements of structure (ibid.). Warning and advertising are both the instance of effectiveness that means the degree to which language users have succeeded in achieving their goals (ibid.). People deliver different views or opinions via ‘linguistics choice’, which is the attitude (ibid.). The medium traditionally has two distinctions: written and spoken. However, with time goes by, it involves the form and style of conveying meaning and information, such as email telephone enquiry (Arndt et al., 2000). Medium affects the public’s thought and use of language. For instance, people just reading through
paper materials like books, magazines, and newspapers, while people can read by watching a video in YouTube (an American Vedic-sharing platform) now. It can not only bring people text information, but also provides people with a better understanding of content with its visual images. A view made by linguists Aitchison (1978), Lyons (1981), and Smith & Wilson (1979) as cited in Aitchison (1981)'s book should be noted that the spoken and written forms of language are separate and relevant systems, and spoken language is regarded as the main form of language. Variety here refers to different forms of language and the context means the situation or environment of language, and there are some fundamental elements, such as form (e.g. written/spoken), style (e.g. formal/informal), social factors (e.g. age, gender, social group) (ibid.). Knowledge of the world indicates that how people use and understand language that is influenced by people’s particular cultural background, learning experience and the way people see the world (ibid.). Besides, users are also an important element in language, what they speak or write depending on how they perceive ‘knowledge of the world’.

Since the 1950s, with the continuous development of the systematic study of language (linguistics), Aitchison (1981) subscribes to language that ‘like everything else, joins in this general flux’ (p. 15). Their book mentions that man of long ago spoke and wrote in so different a manner from contemporary, and previous authors tended to use old English to express their opinions (ibid.). Noticeably, people wrote to the newspaper, published angry articles, and everyone regretted the constant emergence of new words and new pronunciation at that time, which ‘through sheer laziness and sloppiness of mind, we are in danger of losing our past subjective’ and the occurrence of entitled article ‘Polluting our language’ and ‘blind surrender to the momentum or inertia of slovenly and tasteless ignorance and insensitivity’ are cited by Aitchison in 1994 (p. 17). As the research Howard noted in 1980 in the book Words Fail Me that English grammar is going into simpler and rude (Howard, 1980). There are possibilities of human language can be considered. Firstly, the grammar of European languages has tended to decline due to the gradual loss of their old word-ending (Aitchison, 1994), such as word-ending inflectional (is a process of word formation) (‘inflection’, 2020) changes. And then, Language may be slowly developing into a more effective state ‘with existing languages adapting to the needs of the times’ (Aitchison, 1994, p. 19). The author Fairclough (2014) acknowledges contemporary changes influence the role of language in social life. Thus, language and social environment/context interact with each other. Language change includes diachronic and synchronic change. Diachronic change means changing both in terms of historical or chronological perspective, and synchronic change refers “to use theoretical or practical point in time as the basis for analysis” (Arndt et al., 2000, p. 153).

3. Lexical Awareness

Vocabulary knowledge is indispensable in all language skills, such as reading and speaking, listening, and writing. It is one of the constituent elements of language. Vocabulary and language interact with each other. The comment of the relationship between vocabulary and language explored by Nation (1993a) as cited in Schmitt and McCarthy (1997), vocabulary knowledge enables the use of language, the use of language enables the expansion of vocabulary knowledge, world knowledge enables the use of vocabulary knowledge and language, etc. Language is variety, lexis is variety, too. This section deeply illustrates the variety of lexis.

3.1 Medium

The original forms of the medium are spoken and written that have been discussed above. As for lexis, is there any difference between spoken form and written form? The spoken form that produces an English word consists of being able to pronounce the sound in the word and the appropriate syllable stress of the word if it not merely includes one syllable, while one aspect of familiarity with written words is spelling (Nation, 2013). The research of the frequency distribution of core vocabularies may reveal our intuition and awareness about some lexical divergences, whether written or spoken. To understand a word’s spoken form involves being able to recognize it when it is heard, at the other end of the receptive-productive scale, to produce a spoken form to express meaning (Nation, 2013). The corpus-based study made by Nation (2013) reveals that in the top 50 co-occurrence words form 330,000 Cambridge International Corpus written data and 330,000 words of spoken data (CANCODE) indicate the written list including function words (“function words here include all non-lexical, i.e. non-contentful items, such as pronouns, determiners, prepositions, modal verbs, auxiliary verbs, conjunctions, etc.”) (Schmitt & McCarthy, 1997, p. 24). However, the list of spoken words seems to contain many lexical words (means have clear meanings that can describe to someone), such as think, know, and well (ibid.). It is worth to notice that actually function words dominate the top 50 frequencies of the two lists (ibid.). Some lexical words invade high-frequency function word lists and are proved to be elements of interpersonal markers (e.g. you know, I think) or single-word organizational markers (e.g. well, right) (ibid.). The number of vocabularies are seemed as the typical mode of spoken, such as apologies, smooth-overs (e.g. never mind), hedges (e.g. sort of/ kind of), and many other types like the ‘I’ v got’ is emphematical spoken uses (ibid.). In both spoken and written mediums, ‘start’ seems to be the same at home, but in the informal spoken language recorded in CANCODE corpus, ‘begin’ is relatively rare use (Schmitt & McCarthy, 1997). Another example is ‘too’ and ‘also’, which are relatively equal in spoken and written discourse, while ‘also’ appears in written discourse more than half the number of times than it occurs in spoken language (ibid.). According to Schmitt & McCarthy (1997), the top 50 words in written in corpora account for 38.8% of the total texts, and the top 50 words in spoken take up 48.3% of all the text, almost 10% of the total, which confirms that speaking is the primary form of what has been discussed in the second section (reference). Except for the category as spoken and written, new electronic mode as medium arises recently, such telephone, television, computer, film, face to face interaction. As well as "spoken styles intrude on the written medium, such as email is becoming more globally available and central" (Schmitt & McCarthy, 1997, p. 38).

3.2 Users
As for the school (primary school or middle school) students and higher institutions (undergraduate, postgraduate and even PhD) students or scholars, academic words are quite essential to them when they do a presentation, finish assignments, attend the lectures, seminars, and give the speeches, etc. According to Barber (1962) academic vocabulary is generally called useful scientific vocabulary. Usually, academic vocabulary lists gather words such as accumulate, achieve, compound, complex and proportion which commonly appear in academic texts, but not in other contexts (Nation, 2013). Therefore, students in school or college are more likely to use these academic vocabularies, instead of other words.

Moreover, if people as teachers in kindergarten, they would tend to use simple spoken vocabularies that are easily understood by children in the classroom. Because kindergartners can understand oral language far more than they can read, if they are reading, researchers have suggested that vocabulary teaching in preschool and kindergarten be simplified in shared storybook reading (Silverman, 2007). And the view summarised by Wright and Neuman (2014), “There is now overwhelming evidence demonstrating that children’s oral vocabulary development is essential to their long-term reading comprehension.” (P. 330). A wordlist basic vocabulary for beginning reading mention by Johnson (1971) in his article, which at least 50 times were used by kindergartners or first-grade children who took part in the author’s study. Among the word list, the top ranking is the article ‘the’, and ‘a’ ranks the fifth. The word list mainly consists of following categories: preposition (e.g. of, in, with), auxiliary verbs and modal verbs (e.g. was, are, be, could, should, must, may), personal pronouns (e.g. he, me, them), notion verbs (e.g. made, called, get, told), possessive pronoun (e.g. her, his, my), adverbs (always, almost, about), and interrogatives (why, where, what, which, when) within top 200 words.

3.2 Effectiveness

How we choose vocabulary would to a large extent rely on the specific background and context of the communication. There is an example between word choice and medical treatment. “In medicine we use words as tools” in medical education (Wakeman, 2013). When it comes to addiction, patients are described as “substance abusers” (Wakeman, 2013, p. e1). In the forthcoming fifth edition of the Diagnostic and Statistical Manual of Mental Disorders, “substance abuse” and “drug dependence” will be replaced by “substance use disorder” (O’Brien, 2011 as cited in Wakeman, 2013). This is because ‘abuse’ is a term associated with rape, domestic violence and sexual harassment in people’s mind, and the use of such a term ‘substance abuse’ to refer to a chronic, treatable brain disease ignores decades of scientific research suggesting the role of genetics, trauma, and exposure to the neurobiology of disease, and, instead, it tends to a moral view of the individuals ‘choice’ as such (Wakeman, 2013). “Using language that demonstrates an understanding and acceptance of the disease model of addiction will go a long way towards improving the medical treatment of patients struggling with this challenging disease” (Wakeman, 2013, p. e2). Thus, it can be seemed that the word choice or rehearsal language for effectively conveys the meaning and information which is quite significant, and language plays an important role in a number of fields and areas.

3.4 Data-driven learning (DDL) on EFL learners

Data-driven learning studies a large number of English text databases (corpora), containing software programs that are common patterns in authentic language samples (Hadley, 2002). “Data-driven learning (DDL) has become an innovative approach developed from corpus linguistics” (Aşık et al., 2015, p. 87), which plays an important role in the promotion of language pedagogy in vocabulary, because it provides learners with a large number of authentic corpora examples which enable EFL learners to analyze and explore language rules with the help of online corpora and concordance (Aşık et al., 2015). The authors state that statistical analysis of the results shows a positive attitude towards DDL teaching in terms of improving the depth of vocabulary awareness, especially in synonyms and collocations (Aşık et al., 2015). And DDL is considered an effective teaching method for developing vocabulary knowledge compared with traditional teaching methods (Chan & Liou, 2005; Ko, 2005).

Learning to understand a language requires decomposing speech streams into a chunk that reliably mark meaning, which chunk of language could grasp learners’ attention (Schmitt & McCarthy, 1997). “There are limited sets of sounds and written alphabet with vocabulary. “These units occur in more or less predictable sequences” (Schmittn & McCarthy, 1997, p. 122). Using written instances, ‘e’ usually follows ‘th’ rather than ‘x’, ‘the’ is the familiar sequence (ibid.). Likewise, collocations, idioms, and lexicalised phrases are all the actual example of languages occur by chunk which is more easily for the foreign language learner to master and understand the English language.

Actually, in the preceding part of this section, both medium and user influence words choice that come from the DDL, and the vocabulary and the data collected by researchers mostly come from the corpus-based study.

When it comes to where to get the corpora, the common and reliable corpus that can be come into contact with are the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). Both of each can offer at least one hundred million words whatever in spoken and written English, including various disciplines (e.g. education, medical science, law), mediums (e.g. book, journal, magazine), and context (e.g. social context, meeting, radio programmes). COCA has its own interface that brings convince to users.

As for how DDL works in learning vocabulary, here are two typical software as corpus-based tools to explore vocabulary use or language use: Sketch Engine and AntConc (3.5.8 version). Concordance function in both two software could show the collocates of a word on both sides. During the manipulation process of exploring words in the softwares, learners tend to realize that other items (words or phrases) always match the items that they are searching (Hadley, 2002). Taking ‘eaten’ as an example, the most common word following with ‘eaten’ as an object is ‘food’ in BNC by using word sketch function in Sketch Engine (Table 1). And the concordance function shows the collocates of it (Table 2). In table 1, the numbers in the first column refer to the frequency of co-occurrence between ‘eaten’ and its objects in BNC, and the numbers in the second column here mean the typical score (LogDice) that indicates how strong the collocation is.
Table 1w  
The object of ‘eat’ (as a verb) in BNC

Thus, the advantages of the DDL approach for learners are obvious. Learning with the DDL approach, learners could have a visual sensitivity inherent in the word itself. However, due to inadequate conditions, many countries, schools and EFL classrooms do not currently use this approach method for teaching. So far, DDL approach has helped EFL learners to build lexical awareness to some degree.

4. Reflection as a teacher

Context greatly influence people to choose words in language use. It is quite fundamental for EFL learners in learning vocabulary. A good information background and the appropriate situation can provide a wide range of opportunities for students to be exposed to vocabulary. Nation (2013) views it as incidental learning. Actually, learning vocabulary from the contexts can contribute to learning everything about the language. Because students could make conversation/write a story/listen to the video/do an intensive reading by using and recognizing certain words in a particular context. Therefore, setting themes for peculiar words in each lesson will help students to easily understand the words. With the help of the same topic and background, students would guess the meaning of the new words to a great extent, which makes students more impressive to the new words.

In addition, language awareness involves both the language knowledge and the skills possessed by teachers, and teachers’ understanding of their students. Teachers should have a clear mind on the students’ level they teach. So, the appropriate teaching objectives should be set in the appropriate contexts is significant for EFL learners. The degree and level of vocabulary setting should not be too difficult or too simple. Language is variation, and classroom teaching is variation, too. Teachers should adjust their own language expression according to the actual response of students constantly, through the appropriate language medium to teach new vocabulary and other language knowledge.

Consequently, as a teacher, they ought to equip with a good sense of language awareness. This will help to strengthen teachers’ self-confidence and to build the authority of teachers. English teaching in the EFL context is not only to teach the target (English) language culture, but also to teach the language through the culture. The teacher’s language awareness will affect the improvement of the teacher’s language ability and culture ability, and will also affect or contributes to teaching itself to a certain extent.

5. Conclusion

To conclude, being aware of language is significant to language teachers, EFL learners, and even ordinary people. Realizing the variability and flexibility of vocabulary and language would make people use language better and live better.
References: