

PalArch's Journal of Archaeology of Egypt / Egyptology

“YOU KNOW” AS A DISOURSE MARKER IN ENGLISH LEARNERS’ SPOKEN DISCOURSE: UM-SpEAKs

¹Ladyana Rysa Pratiwi, ²Yazid Basthomi, ³Nurenzia Yannuar, ⁴Maria Hidayati

⁵Nabhan Fuad Choiron, ⁶Rofiqoh

^{1,2,3,4,5}, Universitas Negeri Malang, Indonesia

⁶Universitas Tadulako, Indonesia

¹Ladyana Rysa Pratiwi, ²Yazid Basthomi, ³Nurenzia Yannuar, ⁴Maria Hidayati
⁵Nabhan Fuad Choiron, ⁶Rofiqoh, “You Know” As A Discourse Marker In English
Learners’ Spoken Discourse: Um-Speaks– Palarch’s Journal of Archaeology of
Egypt/Egyptology 17(7) (2020). ISSN 1567-214X.

Key words: discourse marker, you know, spoken discourse, UM-SpEAKs

Abstract

Spoken language in English by native speakers is frequently colored with the use the discourse marker *you know*. However, to what extent this holds true with the English spoken language produced by Indonesian learners of English is far from clear. Therefore, this study was set out to investigate this phenomenon by analyzing the spoken corpus of UM-SpEAKs (Universitas Negeri Malang Spoken English in Academic Kontexts). This study analyzes the phenomenon and tries to explains by taking into account the contexts of the use of the discourse marker *you know*. The factors include, inter alia, gender and context of conversation.

INTRODUCTION

In a recent study investigating the use of discourse markers in a multilingual classroom, Martín-Laguna and Alcon-Soler (2018) have shown that discourse markers are one of pragmatic elements that aim to make the conversation or discussion within the classroom more communicative. By adding certain terms of discourse markers, the interlocutor might easily get the point of what the speaker is talking about. Although this may be true, the use of discourse marker terms could be perceived differently by the speaker and interlocutor/s, especially among those who use English as their second or foreign language. Besides, students tend to use discourse markers as verbal aids for task handling and aids supporting reasoning process during task performance (Gánem-Gutiérrez and Roehr, 2011). This finding was concluded based on the number of each certain discourse marker terms used by the students frequently, for example *OK* and *Oh no* which indicate discourse markers as verbal aids for task handling; *so* and *because* which emphasize cause-effect or the function of reasoning process. Even so, we only considered in common discourse markers such as *so*, *because*, *or*, *but*, and *and*. Likewise, a more recent study has revealed more deeply that the unit phrases in discourse markers actually have

different combinations and juxtapositions that make them easier to classify (Cuenca and Crible, 2019). The classification is based on the scope and function in various forms of oral and written communication in everyday life. Again, the data contained common discourse markers that we often encounter such as *and then*, *and so*, and *but actually*. This research was done by examining the function of each various form but it only showed that the primary examined data was a combination of conjunction and connective terms. Another key point, besides the fact that discourse markers are widely recognized as connectors, they also constitute as topic management and speaker shift (Lee-Goldman, 2011). Lee-Goldman's study focuses on principles of conversation and discourse analysis to explain the significance of particular occurrences of a discourse marker *No*. To simplify the classification, the occurrences of the discourse marker were put according to its functions and roles in conversations. Somehow, the previous researchers only showed the importance of this classification arrangement for complex discourse examination markers, and for ongoing discourse representations. Continued to another function of discourse markers, a discourse marker is likely to have more than one function (Lastres-López, 2018). This study prioritized classification in more detail about the function of a common discourse marker that we often hear, but the written definitions were only those that experienced repetition that results in what the researcher wanted to convey was still not clear enough. What is unique about this research is that the discourse markers that are used can also adjust to certain verbs, which makes them involve modals/non-modal utterances. The most significant main discussion in this study was the use of *if*-clause as a discourse marker, so the explanations of other multi-function discourse markers were not explained in detail. It was also stated that this study still needed to be confirmed regarding the types of verb used in the clause.

Additionally, Lastres-López's study has similarity with the study conducted by Haselow (2011) which has confirmed that a discourse marker actually would have different functions when we put it in different orders. This study elaborated the function of final *then* in spoken English as a discourse marker and modal particle. This approximation was also verified by Hancil (2018) in her research entitled *Discourse Coherence and Intersubjectivity: The Development of Final 'but' in Dialogues* which said that "final *but* displays semantic-pragmatic properties of both a discourse marker and a modal particle." Most compelling evidence of Haselow's (2011) study is that it collected corpus data which consist of some situational types of spoken components: private direct conversation, private telephone calls, public classroom lesson, public broadcast discussion and interview, public parliamentary debates, public business transactions, monologues such as spontaneous commentaries on soccer or athletics, unscripted speeches, and broadcast news. Even though the final result of this study is accurate, the corpus that was used was a one-million-word corpus published in 1998. It can be said that this study did not pertain to recent language phenomena which are likely different from the situations nowadays where linguistic aspects or backgrounds can potentially demonstrate new types of language phenomena. Further research has claimed that L2 speakers tend to use *like* as a discourse structurer in conversation and as a focuser device (Magliacane & Howard, 2019). This research uniquely pointed out at the differences and similarities between two learner cohorts from the same background in producing discourse marker *like* and engaged the native speakers as criterion standard. Nevertheless, the data were collected from one-time verbal interviews conducted by the researchers which can be said that it might be less relevant. Moreover, the differences of discourse markers used by these two groups were

not fascinating in describing what aspect that made it different since the respondents were from the same origin and background. In the same way, Buyse (2012) opines that discourse marker *so* which partially has resemblance with *you know* and *I mean* (informality) is a discourse marker that reveals "as a decidedly versatile marker that demonstrates an aptitude to index ideational, interpersonal as well as textual relations." The relevance of the three functions depends on in what way each of them indicates 'resultative' relation, a hypertactic shift, and a point of closure. In the final conclusion, this study conveys that non-native speakers of English (Belgian) who used English as language of instruction in classroom activity frequently used *so*. In particular, it is the comparison between English Linguistics learners and Commercial Science learners which finally identified that they (EL and CL learners) overuse discourse marker *so* than English native speakers do.

From those studies above, it is actually rare to find research which discusses or pertains to discourse marker *you know*. Therefore, we take an initiative to conduct a research project regarding the pragmatics field focusing on discourse marker *you know*. To our knowledge, there has been only one research project discussing discourse markers *you know*, *like*, and *well* (Polat, 2011). In the study entitled *Investigating Acquisition of Discourse Marker through A Developmental Learner Corpus*, Polat confirmed the relationship between discourse markers with pragmatic factors and took developmental learner corpus as its research data. This research was focused on discussing discourse markers used by an immigrant second language learner in the US. The corpus-based research which was conducted within a year compared the researcher's own corpus with SBCSAE or Santa Barbara Corpus of Spoken American English from University of California, Santa Barbara. Concisely, he revealed that the choice of the three discourse marker terms in the highest number of the term *you know* which also pointed out that *you know* itself is used "as a catch-all gap-filler, replacing the disfluency markers and some other discourse markers that are used by native speakers" such as *uh/eh*, *um* (Polat, 2011). Correspondingly, the speaker with his hesitation and disfluencies wants to keep the conversation with the interlocutor going on, so the interlocutor could understand what the speaker wants to convey. Additionally, this research suggested that it would be better to involve several respondents, especially those who have different language backgrounds, educational experience, and proficiency levels.

Overall, based on the explanation above regarding previous studies about discourse markers, the discussion about discourse marker *you know* is an under-researched topic of study. The dearth of discussion regarding discourse marker *you know* leads us to gain more clarity of its use in this present study and the learner corpus comes along in it. Furthermore, the present study intends to provide reasons and meanings behind the usage of *you know* as discourse marker by English learners of Universitas Negeri Malang by analyzing the existing corpus. As such, we referred to UM-SpEAKs (Universitas Negeri Malang Spoken English in Academic Kontexts)—a corpus of spoken data collection based on the recordings which have been gathered from students at the English Department during classroom teaching-learning activities. UM-SpEAKs has several kinds of recording of student performance in spoken classroom learning activities such as panel or group discussion, debate, and group presentation. Earlier studies utilizing the learners' spoken corpus of UM-SpEAKs have been conducted by Hidayati, Choiron, and Basthomi (2019) and Darong, Kadarisman, Basthomi, Suryati, Hidayati, and Niman (2020). As classroom-based analyses, the studies focused on the types of discourse strategies used during students' debate performance in Speaking classes and on questioning practices, respectively. The participants were similar with the

students in this study in the way that they are non-native English speakers and English is applied as the language of instruction. Different from the previous studies, the present study seeks to examine the uses of discourse marker *you know* in the language production of Indonesian learners studying English. The present study is also different from that of Aslan and Şahin (2020) focusing on students' perceptions regarding factors affecting their participation in oral communication course in that the present study has to do with the empirics of students's spoken performances in classrooms rather than their perceptions.

On the basis of the background, the present study discusses *you know* as a discourse marker in English learners' spoken discourse using a corpus study approach to UM-SpEAKs by analyzing the occurrences of *you know* in student spoken performance in classroom. So, the objectives of the study are formulated as follows:

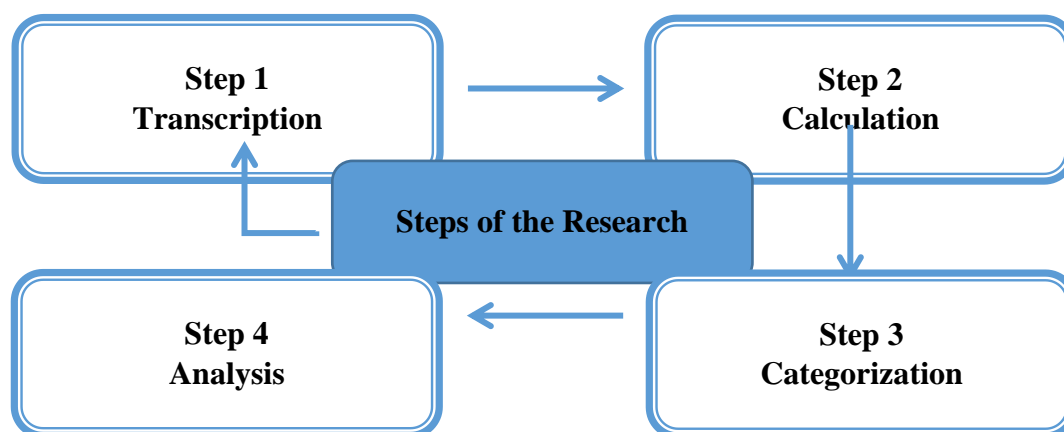
- 1) to identify the meaning and purpose of *you know* used by speakers as a discourse marker in the classroom learning process, and
- 2) to investigate the factors that might make the speakers use *you know*.

METHODOLOGY

Research Design

As adumbrated earlier, this study is corpus-based and focused on the analysis of the meanings and the purpose of the speakers in using *you know* as a discourse marker. Since *you know* has some meanings, this research was aimed to get explanation clearer. This research investigated the use of *you know* by classifying each of it based on its context. The data of this research was the transcription of the recording of classroom learning processes in UM-SpEAKs (Universitas Negeri Malang Spoken English in Academic Kontexts) which consisted of classroom discussions, group discussions, group presentations and debates at the Department of English, Faculty of Letters, Universitas Negeri Malang.

The present study was carried out by taking the following steps:



The first step included the transcription of all the recording of UM-SpEAKs data by using an application, Transcriber AG. The next step was to calculate the occurrences of how many discourse marker *you know* that are used by the speakers. The third step was categorization of the results of calculation based on the context used in the recording. In the final step, we analyzed and examined the results by using an application namely AntConc to know the meaning and purpose of the use of *you know* by examining the contexts of use of the discourse marker *you know*.

Subjects of the Research

The present study was conducted at Universitas Negeri Malang, East Java, Indonesia. Specifically, it took the data of UM-SpEAKs which consisted of the recording of spoken classroom activities including classroom discussion, group discussion, debates, and group presentation. UM-SpEAKs was a corpus of spoken language data collection in which the recordings have been taken from the English Department students during classroom learning activities in their early semesters (AY 2017/2018). Here is the distribution of the subjects recorded in UM-SpEAKs

Table 1 Subjects of the study

No.	Context of Activity	Duration
1	Classroom Discussion	00:35:09
2	Group Discussion	01:01:23
3	Group Presentation	02:17:05
4	Debates	02:19:57
Total		06:13:34

Data Collection

As noted earlier, UM-SpEAKs is a corpus of spoken data collection. The recording was taken from the English Department students during classroom learning activities in 2018. The recording process was done intentionally but not spuriously, so the students proceeded with classroom learning activities as usual. It took students' activities in their spoken performance of classroom and group discussion, presentation, and debates. All of the recordings were compiled in one folder, i.e. spoken data of UM-SpEAKs 2018.

The next step, we transcribed the audio become written data, so the final data was in the form of recording transcriptions of spoken classroom activities based on UM-SpEAKs 2018. The present study extracted the data from UM-SpEAKs and identified the uses of *you know* indicating a variety of purposes and meanings.

Data Analysis

After collecting all of the data from UM-SpEAKs 2018, we sorted out and calculated the data. The data was divided into four types of its use, the classroom discussions, group discussions, group presentations and debates. Then, each of the data was tabulated based on its context to simplify each type.

Step 1 – transcribing all the recording of UM-SpEAKs data by using an Transcriber AG application

We transcribed all the recording of spoken data which consisted of classroom discussions, group discussions, group presentations, and debates. Transcriber AG was used to help us in the process of transcribing. This was done manually to convert audio become written data. This step was aimed to get the frequency of *you know* occurrence.

Step 2 – calculating the occurrence of discourse marker *you know*

After the transcribing process was finished, it continued to this second step—to calculate the occurrence of *you know*. In calculating the result, we used an application namely AntConc.

Step 3 – categorizing the result of calculation based on the context used in the recording

When the data amount had gotten, it was categorized based on its context. Afterwards, we calculated the occurrences of *you know* during the activities (classroom discussions, group discussions, group presentations, and debates).

Step 4 – analyzing and examining the result by using an application namely AntConc to know the meaning and purpose of the use of *you know*

In analyzing and examining the result, we used AntConc again to identify what kind of context the discourse marker *you know* occurred. By using AntConc, it was easier for us to know the frequency and also to find out the full sentences of *you know* used in its context. It was meant to simplify discussing the main topic of the present study.

Research Instruments

In doing this present study, we used two instruments of software applications to facilitate the data analysis in this present study. These instruments were TranscriberAG and AntConc. According to *SOURCE FORGE* website, TranscriberAG is a tool that is “designed for assisting the manual annotation of speech signals. It provides a user-friendly GUI for segmenting long duration speech recordings, transcribing them, labeling speech turns, topic changes and acoustic conditions.”. This transcriber tool was redeveloped by Bertin Technologies and released under GNU GPLv3 License. TranscriberAG itself has 2 main features, each of it has several functions i.e. (i) *format management* to read many audio and video format, handling very long files up to several hours, and synchronized playback; (ii) *transcriptions capabilities* which have several layers of segmentation, support for named entities, easy creation and modification of temporal anchors, and import of old file format.

While AntConc was concordance software developed by Laurence Anthony. It has seven tools and each tool has also its usability i.e. as quoted explanation from *EDUTECH WIKI* (see Table 2).

Table 2 Brief Functions of AntConc

Tools	Functions
Concordance Tool	to show search result in a keyword context.
Concordance Plot Tool	to show search result plotted as a ‘barcode’ format. This allows you to see the position where search results appear in target text.
File View Tool	to show the text of individual files. This allows you to investigate in more detail the results generated in other tools of AntConc.
Clusters/N-Grams	to show how clusters based on the search condition. In effect it summarizes the results generated in the Concordance Tool or Concordance Plot Tool.
Collocates	to shows the collocates of a search term. This allows you to investigate non-sequential patterns in language.
Word List	to count all the words in the corpus and presents them in an ordered list. This allows you to quickly find which words are the most frequent in a corpus.
Keyword List	to show the which words are unusually frequent (or infrequent) in the corpus in comparison with the words in a reference corpus. This allows you to identify characteristic words in the corpus, for example, as part of a genre or ESP study.

To show the practicalities of the two instruments mentioned above in measuring the data, here are the explanation tables of the instruments use:

Table 3 TranscriberAG Use

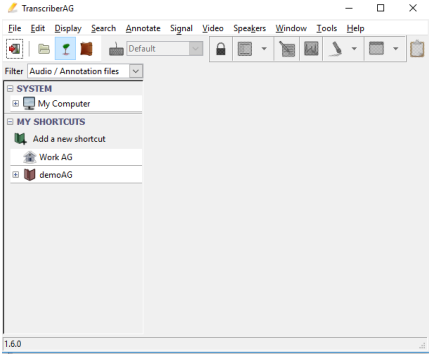
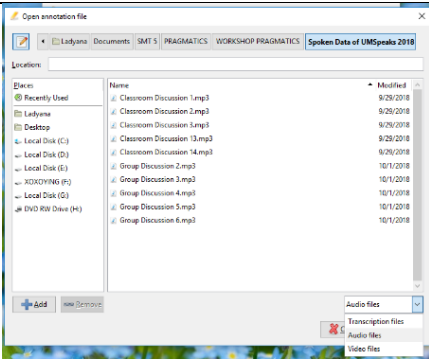
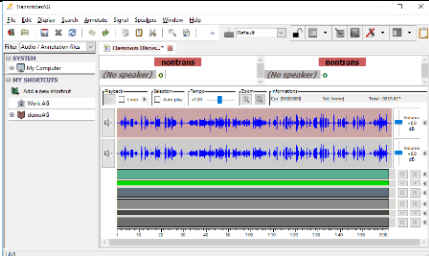
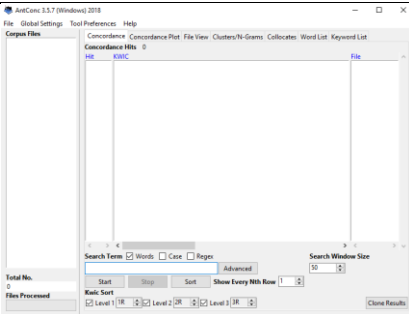
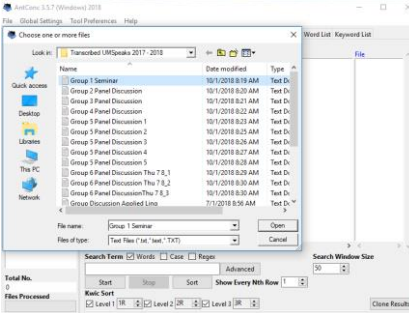
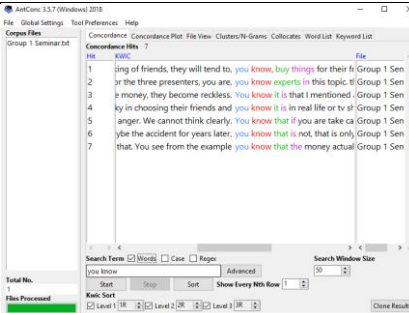
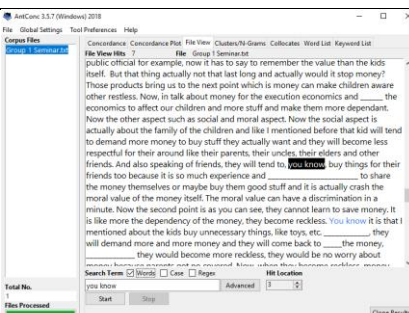
No.	Picture	Explanation
1		<p>This is the main display of TranscriberAG.</p>
2		<p>When we are going to transcribe the audio, we can add audio by choosing <i>file</i> and take the audio that will be transcribed. TranscriberAG can also read the video format.</p>
3		<p>This display occurs after we add audio. While listening to the audio, we could type the transcription of the audio and we could also adjust the audio speed (slow down / speed up).</p>

Table 4 AntConc Use

No.	Picture	Explanation
1		This is the main display of AntConc.
2		We could directly go to <i>file</i> to take the transcribed data. It can take all data in one folder at a time.
3		Type the word or phrase that we want to find out in <i>search term</i> , then it will be evinced along with the full sentence, so we could sort out which sentence that will be examined.
4		To look for the location, we could simply click to the blue words, then it will show the searched world in the whole text.

FINDINGS AND DISCUSSION

This section presents the findings and discussion of the data in view of the objectives of the study aforementioned. The findings show the result of calculating the occurrence of discourse marker *you know* found in the data along with the categorization based on each context (classroom discussion, debate, group discussion, and group presentation). Furthermore, the discussion unfolds factors that possibly trigger speakers to use *you know* by looking at the contexts of the classroom activity and other supported factors such as speakers' background and gender.

Findings

This section examines the first research question regarding the meaning and purpose of *you know* used by speakers as a discourse marker. The findings discussed in this section were the result from the process of calculating and categorizing the occurrence of discourse marker *you know* in the UM-SpEAKs after the transcribing process. There were four contexts of spoken discourse as mentioned in the previous chapter i.e. classroom discussion, group discussion, group presentation, and debates. The findings of calculation in each context was represented into the table, so that there are 4 tables within this section.

Almost all of *you know* which were identified as discourse markers in this present study had certain characteristic which was as the same with Macaulay's study (2000) i.e. it has falling intonation and sounds quite lower than the other utterances. Moreover, its occurrences were unpredictable because sometimes it suddenly appeared in the middle of a sentence that sounds as if it is just said randomly. However, it was not completely difficult to identify *you know* as a discourse marker by listening to the recordings.

Classroom Discussion

Table 5 Occurrence of *you know* in Classroom Discussion

Recording Code	Length	Occurrence of <i>you know</i>
Classroom Discussion 1	00:02:01	-
Classroom Discussion 2	00:03:08	-
Classroom Discussion 3	00:02:36	3
Classroom Discussion 4	00:02:10	-
Classroom Discussion 5	00:03:35	-
Classroom Discussion 6	00:01:37	-
Classroom Discussion 7	00:01:23	-
Classroom Discussion 8	00:03:37	-
Classroom Discussion 9	00:04:37	-
Classroom Discussion 10	00:01:37	-
Classroom Discussion 11	00:01:23	-
Classroom Discussion 12	00:00:36	-
Classroom Discussion 13	00:00:48	-
Classroom Discussion 14	00:01:19	-
Classroom Discussion 15	00:01:52	-
Classroom Discussion 16	00:01:24	-
Classroom Discussion 17	00:01:16	-
Total	00:35:09	3

Inadvertently, there are three *you know* in this session of classroom discussion within the whole duration of 00:35:09, and as can be seen on Table 5 that they occur only in recording code of Classroom Discussion 3. In this recording code, a student discussed whether it is allowed for parents to smack their children or not. To investigate more deeply, *you know* was found within

several sentences, i.e.; The first sentence “*It is they can’t slap in their heads or maybe in their hand you know because there is a fat in that*”, *you know* here is implicitly in accordance with the marker of information structure (Schleef, 2004). It plays a role in emphasizing the information that a person should not give a slap to the head which might be more physically risky, instead they are allowed to hit the hand. This function of *you know* here however are similar with the next two following sentences; 1) “*...and we know that in real life if the children or maybe your sons make some mistake in society, it will be more painful than offend slap you know like people will hate you or maybe will ignore you.*”; 2) “*...after you smack your children, you just angry to your children and make-, make them like mentally-, make them- their mental down you know*”. *You know* as the discourse marker here is used to emphasize or strengthen information that the speakers deliver. In addition, somehow *you know* within those sentences could possibly be omitted without deviating what the speaker intends to talk hence *you know* is recoverable (Fuller, 2003). There will be no difference when *you know* is omitted unless it might only change the syntactical structure.

Group Discussion

Table 6 Occurrence of *you know* in Group Discussion

Recording Code	Length	Occurrence of <i>you know</i>
Group Discussion 1	00:13:19	-
Group Discussion 2	00:14:31	1
Group Discussion 3	00:05:20	-
Group Discussion 4	00:14:27	-
Group Discussion 5	00:06:01	-
Group Discussion 6	00:06:03	-
Group Discussion 6 (video ver.)	00:01:38	-
Total	01:01:23	1

During the group discussion session, there is only one occurrence of *you know* which is written in the following sentence “*...so they can get you know like uh...like uh, imagine uh what will discuss about after this.*”. It can be said that this *you know* is the same as the previous section of classroom discussion that *you know* here could probably be included as the marker of information, yet hesitation / repair marker is the most suitable to define the role of *you know* here. Hesitation / repair marker means that the speaker has doubts to disclose certain information so that *you know* in the above sentence is followed with the other discourse marker of *like* and *uh* to make up or substantiate the information given by the speaker. Moreover, *you know* which commonly appears with *like* indicates a certain vagueness that derives from their original use (Schleef, 2004). In other words, it is assumed that speaker might utter *you know* and *like* with no reason or they just occurred as he/she wants to replace the disfluency markers (Polat, 2011).

Group Presentation

Table 7 Occurrence of *you know* in Group Presentation

Recording Code	Length	Occurrence of <i>you know</i>
Group Presentation 1	00:33:29	6
Group Presentation 2	00:11:30	-
Group Presentation 3	00:15:38	1
Group Presentation 4	00:15:38	-
Group Presentation 5	00:03:16	-

Group Presentation 6	00:19:15	1
Group Presentation 7	00:11:12	-
Group Presentation 8	00:27:03	2
Total	02:17:05	10

In group presentation, high frequent occurrence of *you know* was found. There were some various topics which were discussed including Teaching for Social Changes and Awareness, Online vs Offline Learning, Teacher Proportional Development, Cooperative Learning and Teaching, Technology Information and Communication as the Learning Media, and Teacher and Education in Indonesia. The highest number is on the Group Presentation 1 with 6 occurrences of *you know*. Further elaboration can be seen within the following sentences having *you know* in them:

- 1) “a survey proved that there are-, there is still a lot of *you know* cyber bullying or, or commonly they said sex things towards adults, towards teenager and the age range 13 to 23 in social media”;
- 2) “as *you know* here in Indonesia, of course I believe you-, all of you here use Instagram, right?”;
- 3) “so we are comparing the Indonesia as *you know* how bad social media is but there only 40% of the population there.”;
- 4) “And if *you know* that we are actually in the lowest position...”;
- 5) “...but actually uh *you know* the impact that we have in social media...”;
- 6) “*You know* we only 40% of our populations use social media and now you can see how-, how bad our social media is now.”

You know in the sentences 2, 3, 4, and 6 could possibly constitute no discourse markers because they are mostly part of the syntactic roles especially for the number 4 which denotes syntax of the clause; therefore, it cannot be separated from the former word *if*. Additionally, *you know* in sentence 3 is intended to give background knowledge over the topic discussion. Meanwhile, *you know* in the sentence of number 1 legibly has the role as discourse markers in which they are not related nor involve the syntax of the clause. *You know* in the sentence 1 indicates the mark of boundary which separate the constituents that are ordinarily in adjoining (Macaulay, 2000). Moreover, *you know* as discourse marker in sentence 1 and 5 could potentially have role as the marker of sharing knowledge.

In Group Presentation 3 and 6, *you know* was found in the following sentences; “As *you know* that this presentation will be divided into two-, two sections.”; “In this seminar we will explain one of the platform that maybe many of *you know* uh many platforms that you can use...”. These are *you know* which have no meaning as the discourse marker hence they have the roles as the parts of syntactical form. It is similar to the previous discussion of *you know* in sentences 2,3,4, and 6 of Group Presentation 1.

In Group Presentation 8, there are two *you know* which come to these following sentence; 1) “So if you want to be a teacher you should take PPG or, *you know* PPG right?”; 2) “Cognitive is uh *you know* the intelligent of practicing for the students that we have”. It clearly shows that *you know* in the first sentence cannot be categorized as the discourse marker instead it is used to complete syntactic structure of the sentence. Meanwhile *you know* in number 2 can be categorized as the discourse marker. However, it indicates the speaker attitude which signifies the hesitation marker.

Debates

Table 8 Occurrence of *you know* in Debates

Recording Code	Length	Occurrence of <i>you know</i>
Debate 2	00:23:11	11

Debate 3	00:28:29	26
Debate 6	00:23:50	5
Debate 7	00:30:53	13
Debate 8	00:33:32	9
Total	02:19:57	64

Among all the obtained data in UM-SpEAKs, the debate section shows the highest number of the occurrences of *you know*. Particularly, the discussion consisted of several enticing topics including gender stereotype in society, the disadvantages and the benefits of TV reality show, the need of physical education, child modelling, and the marriage without parents' consent.

In recording code of Debate 2, there are 11 occurrences of *you know* which surely not all of them belong to *you know* as discourse marker. Below are the examples of *you know* that do not play as the discourse marker instead as it is part of syntactic structure.

- 1) “You know that’s the problem or you deny it completely?”
- 2) “...oh a man can also do woman things like you know Ivan Gunawan, a man who’s now you know...”

Meanwhile, *you know* that play as discourse marker are in the following examples of sentence.

- 1) “...we are not trying to eradicate this identity but we concern more about you know the ability to express himself...”
- 2) “...oh a man can also do woman things like you know Ivan Gunawan, a man who’s now you know...”

You know of sentence 1 somehow can be said as the discourse marker since its occurrence does not affect the meaning or purpose of the sentence when it is omitted. In other word, it is recoverable (Fuller, 2003). Moreover, *you know* in sentence 2 can be categorized as the transition marker (Schleef, 2004) where after the speaker gave information about something related to a man could also do what women do, then he/she intended to give example over the main information by giving the example. So that, *you know* is the transition marker between the main information and the supporting information (example). Besides, it can be also as the marker of information which is the sentence almost similar to the sentence example in the section of Group Discussion, “...so they can get you know like uh...like uh, imagine uh what will discuss about after this.”.

However, it was also found that *you know* has ambiguous meanings in which they are assumed to be parts of syntactic construction, yet they sound as if discourse markers as well. The examples are written in the following sentences.

- 1) “It’s like... you know a man can do jobs that most woman do...”
- 2) “Gender filled is like mix of them as well... you know a concept which even a children can can think about it.”

Continued to the next session of Recording Code Debate 3 that there are 26 occurrences of *you know* in which most of them can be categorized as discourse marker. Specifically, most of *you know* in this Debate 3 adjoin another discourse marker i.e. *like*. The elaboration is represented within the example of following sentences.

- 1) “...well, you know the profession TV like the expert or some big business men do their projects and businesses.”
- 2) “...then there is also uh um reality competition like you know the American Idol or The Voice, you know.”
- 3) “Because reality show will just give them a misleading you know like mislead their interpretation about life”

4) “If you are not success to to to lose your weight it will you know like give you more pressure because why can’t I do this and so on.”

In the context of the sentences above, it is as if *you know* play a role in giving pause or a marker when the speakers confirm the information provided by adding other information or example that could potentially support the main information. Therefore, *you know* in close proximity with *like* in those sentences are meant to be a marker that lead to additional information.

Meanwhile in Debate 6, only 5 occurrences of *you know* were identified, which suggests the least number among other data of Debate. Here are the examples of the sentences.

1) “...the activity that build the object of bullying that not not you know not not only students uh can do the the physical education...”

2) “...it will distract their mind and it will you know that uh... distract their mind and it will you know distract them uh then you know, panic.”

3) “So when we do something with our heart you know that will have the excellence result.”

You know in sentence 1 palpably defines speakers’ attitude because there are repetitions of certain words that are assumed to signify speakers’ attitude. Whereas in sentence 2 and 3, the rest three of *you know*, are kind of ambiguous, since they could be parts of syntactical structure and categorized as discourse markers as well.

Hereafter, even though 13 *you know* in recording code of Debate 7 are not all included as discourse marker, there are some unique characteristics found, i.e. *you know* is in a position between two same words / phrase. Here they are given in the following sentences.

1) “It’s also you know it’s also helpful for children and also parents to develop creative minds to make colorful uh costume from you know from waste from uh from used...”

2) “That is you know that is always uh use materials that kids like.”

3) “...when they lose they can you know they can, they can think about how to win fairly.”

4) “...they want to cooperate with the children to you know to... show, to show their ability to do something.”

The same as in Debate 6 that somehow *you know* in this Debate 7 are assumed to show speakers’ attitude over hesitancy in delivering the information. However, they can also be marker which indicates that the speakers need a little more time to emphasize what information speakers wish to reach.

Last but not least, like the data found in the previous sessions that surely not all of *you know* in recording code of Debate 8 can be categorized as a discourse marker, such as *you know* in this first example of sentence “if you are married but don’t you know that one of time to make this marriage certificate?” is obviously part of sentence structure which it can change the meaning of sentence when *you know* is omitted. Whereas *you know* that have role as the discourse markers are given in the following sentences.

1) “This kind of thing cannot be done only by having uh you know just the both.”

2) “...uh in number second justification that the parents will be able to accommodate those problems or you know, try to fix their relation.”

3) “And it helps you to you know to be your... to be your, to be your help in the first place.”

Discussion

This section discusses specifically the factors that might make the speakers use *you know*. The factors are elaborated from the viewpoints of the contexts of conversations, speakers’ background, and gender.

Contexts of the Conversation

There is a significant difference in the occurrences number of the discourse marker *you know*, which the first is clearly written because of the difference of the lengths in the recording data. While putting the length factor aside, every context of the conversation itself assuredly plays a role in influencing the speakers to use discourse marker *you know*. As illustrated in the table 3.2.1, the order of the use of discourse marker from the smallest to the largest number is Group Discussion with the occurrence number of 1, Classroom Discussion and Group Presentation having the same number of 3, and Debate with the highest number of 62.

Table 9 The use of discourse marker *you know*

Contexts	Number of <i>you know</i> as DM
Group Discussion	1
Classroom Discussion	3
Group Presentation	3
Debate	56
Total	63

The first is Group Discussion. The role element in this context of Group Discussion is generally like the other group discussion, that is led by one moderator and there are also some panelists. In addition to simple discussion topics, it turned out that the discussion panelists only delivered brief elucidations over the questions previously given by the lecturer. In other words, they (speakers/panelists) in most case only answered the questions without any deeper discussion such as exchanging ideas and opinions between panelist members in their group. The panelists delivered each of their elucidation that they had previously compiled, so that their utterances seemed fluent and smooth, and eventually used less discourse marker. Reflecting on Haselow's (2019) explanation that Discourse Marker is a lexical expression that can be formal and functional regardless of originating and arising from natural speech. Meanwhile, as has been explained that the discussion system in the context of this Group Discussion is that speakers have prepared elucidations or answers to questions given by the lecturer before they started the recording, so that their conversation however sounded less natural.

The second, the context of Classroom Discussion obtained 3 occurrences of discourse marker *you know*. Unlike the Group Discussion speech delivery system, that in the context of this Classroom Discussion, instead of a panelist or speaker delivering their speech only in the scope of small group, they delivered it in front of the class. Overall recorded data, each recording consisted of only one speaker and there is no moderator in the classroom discussion. That is to say, speech delivery system of the Classroom Discussion had deeper discussions including agreeing and rebutting the opinions or explanation of other speakers and also providing alternative suggestions on topics discussion given by the lecturer. Thus, the conversations happened as the semi-natural which made the use of discourse marker *you know* was higher than the context of Group Discussion.

The third is the context of Group Presentation. The number of occurrences of discourse marker *you know* is the same as Classroom Discussion, which is only 3. The difference is, in the context of group presentations, discourse marker *you know* was said by two speakers. Besides, the context of group presentation is

surely dissimilar from the context of classroom discussion, in which each recording data of group presentation consisted of a moderator, 3 to 4 presenters, and the rest of students in the classroom also participated as the audience which later asked question after the group presentation.

The last one, in the context of the debate, has the highest number of discourse marker *you know* occurrences i.e. 63. Apart from the length duration factor which also has the longest recorded data among other contexts, surely the context of this debate has a more specific factor i.e. the topics of debate that are heavier than in other contexts, and the debate system itself. Debate is a speaking activity consisting of 2 teams which requires speakers in each team to critically draw up counter-arguments in order to convince and defeat the opposing team argument (Hassan, 2020). Consequently, speakers tend to speak naturally without depending on the text that might have been prepared previously such as in the context of Group Discussion and Classroom Discussion. In this context of debate, speakers listened and delivered the elaborations related to motion topics for longer time durations, so that the responses in the form of rebuttals or agreements also took long times. When the speakers delivered their ideas over the motion topic extensively, they were required to make their interlocutors understand and follow the speakers' speech flow. Here is where discourse markers have role to help speakers managing the flow of conversation. Therefore, there are more use of discourse marker *you know* in this debate context which are relatively not in accordance to the grammatical function and semantic structure (Popescu-Belis and Zufferey, 2010), such as in the following example of sentences:

- "...that men should do manly things like uh you know being an architect and that woman cannot do..."
- "They can you know easily the... the... what is it?"
- "So it will distract their mind and it will you know that uh... distract their mind and it will you know distract them uh then you know panic."
- "I believe that this 21st century uh scientist in this world already provide you know make- makeups that already safe for kids."
- "...they make claim that the idea of psychology will you know it changes the idea of parents."

Speakers' Background

Furthermore, in this second factor, we would focus on speakers' background in using discourse marker of *you know*, in which discourse marker *you know* itself could be said that its frequency use might tend to be less than the other popular discourse markers commonly used in English speech such as *and*, *but*, *so*. On this speakers' background, there are two main constituents that might potentially influence the use of discourse marker *you know*, i.e. learners' pragmatic awareness, teachers' practice and sociolinguistics context (Martin-Laguna and Alcon-Soler, 2018). However, this present study took factor of sociolinguistics context since English is not the main language used by speakers in daily communication, yet they only use English during teaching and learning activities.

In this case, all students are not English native speakers, therefore, they tend to speak Bahasa Indonesia and Javanese language in their daily lives to communicate with each other. It has been known that Javanese language is the language most widely spoken in the Java island of Indonesia, especially in East Java province. Moreover, the language itself has various kinds of accents and dialects that differ from a certain city or region to others. Only a few areas of East Java use other than Javanese as the language of daily communication. Malang City, which is the place where the sample of this study was obtained,

also uses Javanese as its mother tongue and also daily language. Even though there are some students who do not come from Java, they could not help to mingle with their friends who are mostly Javanese. Almost all students communicate in their daily lives using Javanese and some also use Bahasa Indonesia. Meanwhile, discourse marker *you know*, if it is interpreted into Bahasa Indonesia become '*kamu tahu*', '*kamu mengerti*', and into Javanese also has several interpretations since Javanese itself has several dialects become '*kon ngerti*', '*awakmu ngerti*', '*sampeyan ngerti*', '*panjenengan sumerep*', '*panjenengan ngertos*', and several other variations that cannot all be mentioned. Those interpretations mean as if the speaker expects or guesses the interlocutor(s) already know what certain information that the main speaker delivers.

Due to the habit of students using Javanese and Indonesian as their daily language of communication, here we argue that it would be possible that they interpret their daily speaking attitudes during their learning sessions where English is the main language of instruction during the class activities. Let's have a look, for example, at the following sentences:

Origin version (in English): "*So when we do something with our heart you know that will have the excellence result*"

Translated version (in Bahasa Indonesia): "*Jadi ketika kita melakukan sesuatu dengan tulus kalian tahu itu akan membuahkan hasil yang bagus*"

Translated version (in Javanese language): "*Dadi lek pas awake dewe ngelakoni prekara kanthi ati sampeyan ngerti iku bakal menehi kasil sing apik*"

Those two translated versions are the standard versions, meanwhile students tend to speak in informal way. Here are the examples:

In Bahasa Indonesia: "*Jadi pas kita ngelakuin suatu hal pake hati kalian ngerti kan itu bakal ngasih hasil yang bagus*".

In Javanese language: "*Dadi lek awake dhewe ngelakoni hal sing tulus tekan ati kon ngerti iku bakal menehi hasil sing apik*".

In this matter, we strongly construe that students' social environment along with its sociolinguistics context might potentially influence them to use *you know* as discourse marker during their lecture performances. Furthermore, factor of sociolinguistics context here is the same as where the speakers come from and live in it which influence their way of speaking (language). For the most part, the use of discourse marker *you know* in this case is roughly affected by everyday social life and speakers' background including their mother tongue (Bahasa Indonesia and Javanese language).

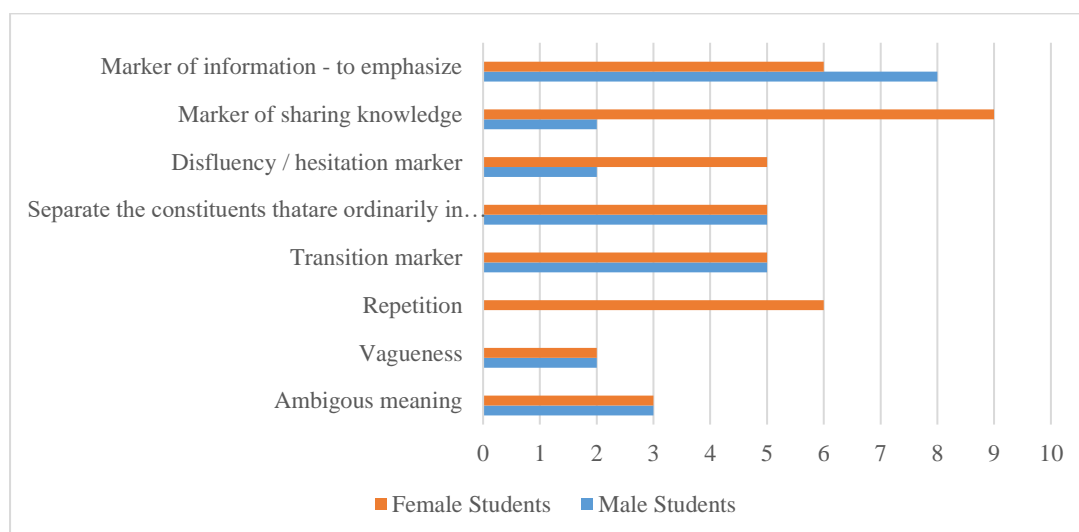
Gender

The last factor is gender. After sorting it out, it turned out that there were quite a lot differences in the number of occurrences of discourse marker *you know* between male and female students. Matei (2011) argues that the gender factor is the most important thing that influences the use of discourse markers in a conversation, which in her research results, she found that female students use more discourse markers than male. In the same way, the result of this present study found that male students employed discourse marker *you know* less than female students i.e. with a total number of 15. Meanwhile, female students used discourse marker *you know* two times more than the total male students had, i.e. 39. The result was represented in Table 10.

Table 10. Different use of discourse marker *you know* between male and female students

Contexts of Conversation	Number of <i>you know</i> as DM said by:	
	Male students	Female students
Classroom Discussion	3	-
Group Discussion	-	1
Group Presentation	1	2
Debate	20	36
Total	24	39

These results are almost similar to that of Macaulay (2000) who found that both female more frequently used discourse markers *you know* than male did and the different number was also almost a half of it. However, when the differences were classified based on the role of discourse marker *you know* and the number of its use by male and female students, there were also particular differences in the way male and female students used the discourse marker *you know*. As Haselow (2019) and Fischer (2012) agreed that most likely a discourse marker could have several functions, so that a single discourse marker *you know* could be put on one or more categories of its use. The result of the classification is represented in Figure 2.

**Figure 2 The differences of *you know* roles used by male and female students**

From the result above, it can be concluded that gender differences trigger students to use discourse marker *you know* at some points in a different way. In this case, female students dominate the use of *you know* as the marker of sharing knowledge, disfluency or hesitation, and repetition. Meanwhile, male students tend to use discourse marker *you know* to emphasize the given information, and they even do not put it as repetition at all. Apart from this difference of discourse marker *you know* role, as a matter of fact, the use of discourse marker *you know* by male and female students here mostly they take it to serve the same purpose, yet the difference is in the way they express it, so that it presumably has different role when it is used by male and female students. In other words, it cannot be clearly said that there are specific roles of

discourse marker *you know* when male or female employed it. Hence, it is in the difference of the way they express it and how the interlocutor or other audience consider its meaning. Besides, there are half of the 8 functions of *you know* in which the frequency of the use between male and female students have the same number, i.e. to separate the constituents that are ordinarily in adjoining, transition marker, vagueness, and ambiguous meaning. These results prove that although there are differences in the frequency of using of discourse marker *you know*, in fact both male and female students are still in certain contexts using it without any differences. Thus, I argue that gender factor might determine the use of discourse marker *you know*, yet it is not predominantly influence its occurrences and roles.

CONCLUSIONS

This section presents conclusions drawn from findings and discussion in the previous chapter. The conclusions succinctly reiterate the answers to the research questions. Referring to the research questions about the meanings and factors of English learners using *you know* as a discourse marker during the classroom learning process, we have found several important points on those questions. Firstly, this present study found that discourse marker *you know* has broadly various meanings and purposes when they are viewed based on the context or situation of its use, especially during students' performances of spoken discourse including classroom discussions, group discussions, group presentations, and debates. Based on the data obtained, there were 63 occurrences of *you know* as discourse marker within the duration of 06:13:34 in student performances of spoken discourse. The 63 occurrences of *you know* as the discourse marker certainly had different characteristics in each of the contexts.

In the context of classroom discussion with the occurrence number of 3, *you know* as discourse marker that had been used were meant to be the marker of information. Similarly, the use of discourse marker *you know* in the context of group discussion with the occurrence number of 1, likewise had a role as marker of information, yet it could be also considered as the hesitation marker. Meanwhile in the context of Group Discussion with the occurrence number of 3, they had two functions including the marker of sharing knowledge and the mark of boundary which separated the constituents that were ordinarily in adjoining. Then, in the context of debate with the occurrence number of 56, discourse marker of *you knows* here had more meanings than the previous contexts along with unique characteristics other than as marker of information and sharing knowledge i.e. some had ambiguous meaning (could be identified as a discourse marker and part of syntactic structure) and vagueness. Moreover, in certain moments, it was also found *you knows* that adjoin other discourse marker, *like* which meant to be the marker that lead to certain additional information.

Furthermore, there were *you knows* that meant the marker of hesitancy which were performed of repetition words that were in front of or next to the discourse marker *you know*, and there were also *you know* which were flanked by two similar words or phrase. However, those last two characteristics could be categorized as hesitancy marker which showed speakers attitude.

Hereinafter, there were three factors that triggered speakers to use *you know* as a discourse marker during their performance of spoken discourse i.e. context of the spoken discourse, speaker's background, and gender. Yet, from the further analysis of gender factor, actually it did not really affect the use of the discourse marker *you know* because the analysis data obtained showed almost the same results between male and female students in using *you know*. Besides, the factor of speaker's background was the main factor that triggered the use of

you know as a discourse marker, which was closely related to sociolinguistic factors. In this case, it was presumably that speakers interpreted their daily speaking attitude into their lecture sessions where English was the main language of instruction during the class. Additionally, the context of the conversation was a second factor that was no less important which consisted of classroom discussions, group discussions, group presentations, and debates. In this factor, the topic and system or model of the conversation also highly influenced the use of discourse marker *you know*. For the most part, factors that potentially triggered students to use *you know* as a discourse marker during their performances of spoken discourse were sociolinguistic context and the context of conversation itself.

REFERENCES

- Aslan, R. and Şahin, M. (2020). 'I feel like I go blank': Identifying the factors affecting classroom participation in an oral communication course. *TEFLIN Journal*, 31(1), 19-43. DOI: <http://dx.doi.org/10.15639/teflinjournal.v31i1/19-43>
- Bolden, G. B. (2015). Discourse markers. *The International Encyclopedia of Language and Social Interaction*, 1-7. doi:10.1002/9781118611463.wbielsi031
- Buyse, L. (2012). So as a multifunctional discourse marker in native and learner speech. *Journal of Pragmatics*, 44(13), 1764–1782. doi:10.1016/j.pragma.2012.08.012
- Cornish, F. (2006). Understanding spoken discourse. *Encyclopedia of Language & Linguistics*, 227–230. doi:10.1016/b0-08-044854-2/04296-6
- Cuenca, M. J., & Crible, L. (2019). Co-occurrence of discourse markers in English: From juxtaposition to composition. *Journal of Pragmatics*, 140, 171–184. doi:10.1016/j.pragma.2018.12.001
- Darong, H.C., Kadarisman, A.E., Basthomi, Y., Suryati, N., Hidayati, M., Niman, E.M. (2020). What Aspects of Questions Do Teachers Give Attention To? *International Journal of Innovation, Creativity, and Change*, 10(11): 191-208.
- Diskin, C. (2017). The use of the discourse-pragmatic marker “like” by native and non-native speakers of English in Ireland. *Journal of Pragmatics*, 120, 144–57. doi:10.1016/j.pragma.2017.08.004
- EDUTECH WIKI: AntConc. <http://edutechwiki.unige.ch/en/AntConc>
- Fischer, K. (2012). Discourse markers. *The Encyclopedia of Applied Linguistics*. doi:10.1002/9781405198431.wbeal0330
- Fox Tree, J. E. (2007). Folk notions of um and uh, you know, and like. Text & Talk – An Interdisciplinary Journal of Language. *Discourse Communication Studies*, 27(3), 297–314. doi:10.1515/text.2007.012
- Fuller, J. M. (2003). The influence of speaker roles on discourse marker use. *Journal of Pragmatics*, 35(1), 23–45. doi:10.1016/s0378-2166(02)00065-6
- Gánem-Gutiérrez, G. A., and Roehr, K. (2011). Use of L1, metalanguage, and discourse markers: L2 learners’ regulation during individual task performance. *International Journal of Applied Linguistics Vol. 21 _ No. 3*

- Hancil, S. (2018). Discourse coherence and intersubjectivity: The development of final but in dialogues. *Language Sciences*, 68, 78–93. doi:10.1016/j.langsci.2017.12.002
- Haselow, A. (2011). Discourse marker and modal particle: The functions of utterance final then in spoken English. *Journal of Pragmatics*, 43(14), 3603–3623. doi:10.1016/j.pragma.2011.09.002
- Hassan, S. (2020). *The Art of debate*. doi: 10.13140/RG.2.2.12397.33765
- Hellermann, J., & Vergun, A. (2007). Language which is not taught: The discourse marker use of beginning adult learners of English. *Journal of Pragmatics*, 39(1), 157–179. doi:10.1016/j.pragma.2006.04.008
- Hidayati, M., Chiron, N.F., & Basthomi, Y. (2019). Students' discourse strategies in a classroom debate performance. *KnE Social Sciences*, 3(10), 460-467. <https://doi.org/10.18502/kss.v3i10.3934>
- Lastres-López, C. (2018). If -insubordination in spoken British English: Syntactic and pragmatic properties. *Language Sciences*, 66, 42 - 59. doi:10.1016/j.langsci.2017.12.009
- Lee-Goldman, R. (2011). No as a discourse marker. *Journal of Pragmatics*, 43(10), 2627–2649. doi:10.1016/j.pragma.2011.03.011
- Liu, B. (2013). Effect of first language on the use of English discourse markers by L1 Chinese speakers of English. *Journal of Pragmatics*, 45(1), 149–172. doi:10.1016/j.pragma.2012.11.002
- Macaulay, R. (2002). You know, it depends. *Journal of Pragmatics*, 34(6), 749-767. doi:10.1016/s0378-2166(01)00005-4
- Magliacane, A., & Howard, M. (2019). The role of learner status in the acquisition of pragmatic markers during study abroad: The use of “like” in L2 English. *Journal of Pragmatics*, 146, 72-86. doi:10.1016/j.pragma.2019.01.026
- Martín-Laguna, S., & Alcón-Soler, E. (2018). Development of discourse-pragmatic markers in a multilingual classroom: A mixed method research approach. *System*, 75, 68–80. doi:10.1016/j.system.2018.03.009
- Matei, M.G. (2011). The influence of age and gender on the selection of discourse markers in casual conversations. *Bulletin of the Transilvania University of Braşov, Series IV: Philology & Cultural Studies*, 1, 213-220.
- O'Connel, D. C., & Kowal, S. (2009). *Transcription for spoken discourse. The Pragmatics of Interaction*.
- Polat, B. (2011). Investigating acquisition of discourse markers through a developmental learner corpus. *Journal of Pragmatics*, 43(15), 3745–3756. doi:10.1016/j.pragma.2011.09.009
- Popescu-Belis, A., & Zufferey, S. (2011). Automatic identification of discourse markers in dialogues: An in-depth study of like and well. *Computer, Speech & Language*, 25(3), 499–518. doi:10.1016/j.csl.2010.12.001

Schleef, E. (2004). Gender, power, discipline, and context: On the sociolinguistic variation of okay, right, like, and you know in English academic discourse. *Texas Linguistic Forum* 48: 177-186.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.471.7751&rep=rep1&type=pdf>

SOURCEFORGE. <https://sourceforge.net/projects/transag/>

The Glossary of Education Reform. (2013). <https://www.edglossary.org/english-language-learner/>

TranscriberAG: *a tool for segmenting, labeling and transcribing speech*.
<http://transag.sourceforge.net/index.php?content=features>

UK Essays: Provider of Free Study Resources. (2003).
<http://www.ukessays.com/essays/english-language/spoken-discourse.php>