

---

## Investigating Pronouns: The Use of “we” in Secondary School Students’ Essay writing

Moustafa Mohammed Shalabi, University Utara Malaysia, Malaysia, Email: [mostafa.shalaby1970@gmail.com](mailto:mostafa.shalaby1970@gmail.com)

Nagi Saad Fathalla Nagi, Al-Baha University, Email: [saadnagi8@gmail.com](mailto:saadnagi8@gmail.com)

Mahmmed Said Negm, Prof. of linguistics, Tanta University, Egypt, Email: [ngm\\_msaid@hotmail.com](mailto:ngm_msaid@hotmail.com)

---

### ARTICLE INFO

Received: December 24, 2019

Accepted: January 26, 2020

Published: January 31, 2020

Volume: 1

Issue: 2

---

### KEYWORDS

*Pronouns, Computer-assisted Corpus Analysis (CACA), Secondary School Students’ Essays, Part-of-speech (POS), CLAWS Tagger, Corpus*

### ABSTRACT

The present study is an attempt to analyze the linguistic constituent; pronoun, “we” as stated in the essays written by third year secondary students of Arab Republic of Egypt in two Secondary schools. As the study under investigation depends so much on a corpus and descriptive analysis, it explores the L2 output of learners in their writing of essay compositions context. It provides examples of pronouns employed by the students. This particular study was aimed at investigating pronouns’ frequencies, the singular–Pronouns, and the Plural Pronouns of the written texts, highlighting the students’ understanding of making use of this linguistic constituent when they write. A corpus-based learner utilized a corpus investigation that has tagging and a frequency of part of speech (POS) of examination employing a software known as concordance will be the scheme of analysis. The results depicted an over-use of the 1st person plural as a subjective personal pronoun (we). The analysis discussion of the pedagogical contribution will be introduced.

---

## 1. INTRODUCTION

To assess, evaluate and access linguistic units in the language of a student has been the main idea of many investigations lately. The collection of the representative learner corpus provides accommodations of this investigation. The expression "representative learner corpus" employed in the study point to a corpus which has exactly been collected to complete a particular investigation, and this does not mean using it for any supplementary reasons. The objective of this study is to present an investigation of the Pronouns-forms existed in the collection of a representative learner corpus recognized as Secondary Schools English Language Learner Corpus (SSELCL).

The current study will initially explain the methodology participatory in the part-of-speech tagging (POS) tailed by an examination of the associated linguistic constituent of SSELCL. Lastly, a discussion of the pedagogical input and the study limitations will be presented. The current study seeks to find an answer to these questions:

1. What are the various forms of Pronouns utilized in the collected corpus of SSELCL?
2. What is the possible distributional patterns of the Pronouns forms observed in the SSELCL?

There are many researches that have described the findings of corpus-based researches that were directed so as to investigate the syntactic grammar and the morphology of several learner languages. The morphological examination results are employed to upgrade the systems of translation of one language to another. The studies that examined learners’ corpora which have a collection of the language learners’ translation to English from their mother tongue are few such as those by (Popović, Ney, & Vilar, 2005) who examined the Serbian-English translation, (Lee, 2004) the Arabic-English translation, and Goldwater and (Popović, Ney, & Gispert, 2006, Goldwater & McClosky, 2005). In the analysis of Aijmer, (2002), She explored learners’ corpora so as to compare modal demonstrated by native speakers and Swedish learners of the English language. The results point to the using of modals is very high by the learners of Swedish comparing to the use of modals used by the native speakers of English.

Zhang, (2008) examines the reasons which effect of utilizing and deleting of ‘be’ verb in two dissimilar learner corpora. By investigating the corpus, the analysis of concordance introduces an overall view of ‘be’ distributional patterns in Chinese-English interlanguage grammars. Making notes on the corpora, the tagging system employed by Yanyan that established by Prof Hung of Hong Kong, Baptist University. An investigation locally directed by Bee& Chan, (2004) inspected the inflectional ‘-s’ morpheme acquisition (genitive marker, the marker of plural, and the marker of third person singular in the present tense).

The respondents are 18 learners of ESL taken from two Chinese primary schools which are in Malaysia. The respondents are given a picture and they are required to give a description to what they have seen in the photos in

English. The respondents' oral explanations formed a spoken corpus after recording and transcribing them. The results indicated that the learners showed a distinctive accurateness morphemes order, and erraticism side by side with L2 utterances overgeneralizations, though the '-s' morpheme acquisition was staged and systematic.

Lorenz (1998;1999) explains adjective strong increase through investigations which concentrated on specific linguistic characteristics. Also, Aronsson (2001; 2003) illustrates on clefts and extra position. Granger and Tyson (1996) conduct investigations which concentrated on specific linguistic characteristics. Granger and Tyson (1996) conduct on connectors and participle clauses are elucidated by Granger (1997).

Cook, (1994) states that the main aspect investigated is the linguistics specialization, namely, what is known about the language by the learner, and linguistics performance, that refers to the capacity of the student to utilize the language.

In accordance with investigation done to study and analyze Noun; the linguistic constituent, as shown in the essays that the students of business and management have written in two higher Malaysian learning organizations in Journal writing, Essay writing and Media invitation to discover and examine the L2 output learner in the context of business, providing examples of NOUN employed. The findings depicted an over usage of the form of singular noun (Kaur & Shamsudin, 2011; Kaur et al., 2014).

Learner corpus should make an exceptionally beneficial application to examine language studying operations and also to increase the provide language studying techniques. Leech (1998) states that a learner corpus is a computerized textual database that the learners of foreign language produce. Similarly, Granger (1998), points out that a learner corpus is made basically via collecting impartial data that can support and help giving a description to learners' language. Researches of a learner corpus contains the tagging of POS, error tagging, tagging of discoursal, tagging of morpho-syntactic and parsing.

Even though limited, learner corpus expansion is also acknowledged consideration in Malaysia. The EMAS corpus, developed by investigators of university of Putra Malaysia, comprises of unedited and untagged written data done by about 800 secondary and primary schools learners. An ongoing project, the Malaysian Corpus of English (MEC) being made by a group of investigators from Malaysian university, is collecting a corpus in the essays form that is written by undergraduates of Malaysian university.

Correspondingly, the main goal of the present research seeks to define and disclose L2 learners' aptitude to utilize English language in Essay Writing, via a frequency examination of the written learners' output. The concentration-is to decide, state the usage of the linguistics constituent, Pronoun. The employed approach is presented in Sarimah Shamsudin and Manvender Kaur (2010).

## **2. INSTRUMENTATION**

The corpus-based learner analysis comprises a Part of Speech tagging that employed a tagger which is online available. Based on the examined size of the corpus, the Part of Speech tagger could be utilized online [ucrel.lancs.ac.uk/claws/trial](http://ucrel.lancs.ac.uk/claws/trial) or for a little payment. Regarding the collected size of SSELc is small, the current study performed the examination by tagging software. The frequencies of the related linguistic constitutions are computed using AntConcordance3.4.4w (Windows) 2014, a concordance software that is available freely and the researchers can be download it online.

The corpus of SSELc is gathered utilizing the rich format of text. Then it is tagged for Part of Speech, employing the Constituent Likelihood Automatic Word tagging System (CLAWS) tagger (the CLAWS7 Tag set). In the early 1980s, the CLAWS tagger is developed by the UCREL from the University of Lancaster and has been improved continuously. CLAWS 4 is utilized to tag the British National Corpus that has 100 million words (R Garside, 1987; R Garside, Fligelstone, & Botley, 1997; Roger Garside & Smith, 1997; Leech, Garside, & Bryant, 1994; Yusuf & Qismullah Yusuf, 2009). Part of Speech tagging is also known as grammatical tagging, is one of the most popular form of corpus annotation. The tagger of CLAWS has consistently achieved about 96-97% accuracy having a mistake rate of about 1.5%. The Part of Speech tagging is supplied by a list of coding for the CLAWS (7) Tag set. The tagger that supports three modes of tagging as well; vertical horizontal, and html is accessed by the internet at the trial service of CLAWS WWW. It is useful to obtain a general descriptive view of the linguistic constituents of the individual text by the horizontal tagging.

Table (1) shows an instance of the horizontally tagged text.

299 words tagged: Tagset: c7 Output style: Horizontal  
 No\_PN121 one\_PN122 can\_VM deny\_VVI that\_DD1 being\_VBG with\_IW your\_APPGE  
 family\_NN1 is\_VBZ the\_AT best\_JJT thing\_NN1 That\_CST makes\_VVZ you\_PPY  
 happy\_JJ .\_.  
 Playing\_VVG with\_IW your\_APPGE brothers\_NN2 ,\_, talking\_VVG to\_II your\_APPGE  
 sisters\_NN2 and\_CC helping\_VVG your\_APPGE mother\_NN1 with\_IW cooking\_NN1  
 are\_VBR all\_DB simple\_JJ things\_NN2 that\_CST have\_VH0 the\_AT ability\_NN1 to\_TO  
 make\_VVI you\_PPY happy\_JJ for\_RR21 long\_RR22 time\_NNT1 .\_.  
 Yesterday\_RT ,\_, was\_VBDZ a\_AT1 very\_RG special\_JJ day\_NNT1 for\_IF me\_PPI01  
 .\_.  
 It\_PPH1 was\_VBDZ my\_APPGE younger\_JJR sister\_NN1 's\_GE birthday.My\_NNU  
 father\_NN1 decided\_VVD to\_TO celebrate\_VVI it\_PPH1 in\_II Alex\_NP1 -\_- the\_AT  
 town\_NN1 which\_DDQ we\_PPIS2 love\_VV0 most\_RRT .\_.  
 My\_APPGE father\_NN1 was\_VBDZ born\_VVN in\_II Alex\_NP1 and\_CC all\_DB our\_APPGE  
 relatives\_NN2 are\_VBR there\_RL .\_.  
 First\_MD ,\_, we\_PPIS2 went\_VVD to\_II the\_AT beach\_NN1 where\_CS we\_PPIS2  
 played\_VVD all\_DB together\_RL .\_.  
 Next\_MD after\_CS we\_PPIS2 had\_VHD finished\_VVN our\_APPGE lunch\_NN1 ,\_,  
 we\_PPIS2 went\_VVD to\_II the\_AT Cinema\_NN1 to\_TO watch\_VVI a\_AT1 funny\_JJ  
 film\_NN1 .\_.  
 My\_APPGE sister\_NN1 love\_VV0 this\_DD1 kind\_NN1 of\_IO films\_NN2 so\_RG much\_DA1  
 .\_.  
 .\_.

It is essential to tag the corpus horizontally so as to ease the frequency calculation utilizing the software of concordance. The horizontally tagged version of the corpus is coded individually, and is downloaded with the purpose of the frequency examination.

### 3. METHOD

For the purpose of scrutinizing the linguistic constitutions of the output of written Essay-compositions, the current study collected a learner corpus that comprises of the writings of two secondary schools in Egypt; one being a secondary school for girls while the other is a secondary school for boys. The text of the corpus is coded individually, according to the written category forming essay writing.

The codes used in this study 3SEF1 points to third year Essay written by female respondent one, 3SEF2 points to third year Essay written by female respondent two and 3SEF3 points to third year Essay written by female respondent three. 3SEM1 points to third year Essay written by male respondent one, 3SEM2 points to third year Essay written by male respondent two and 3SEM3 points to third year Essay written by male respondent three.

The compiled corpus is called SSELCS Secondary School English Linguistic Corpus SSELCS for brevity, substitute the essay writing composition of corpus context of t. A significant notice to be taken in to consideration before going further is that in the present study, only written texts are employed and accumulated into a learner corpus because of the supposition which real writing represents learners' grammatical competency carefully more than speech. The categories observed in the present corpus of the current study are original written materials from the assignments of essay writing, media invitations and the activities of journal writing that the students have written registered in the department of business and management in the institutions of higher learning.

### 4. RESULTS

The investigation of frequency is done utilizing the AntConc 3.4.4 w (Windows) software. AntConc3.4.4w (Windows) is developed by Laurence Anthony (2014), Faculty of Engineering and Science, University of Waseda, Japan. In addition to doing the functions of plotting and concordance, the AntConc3.4.4w (Windows) is significant to compute and collocate tags frequency.

The computed frequency by the concordance is analyzed regarding the various tags allocated to the various linguistics constituents exactly as employed in the CLAWS tagger (see codes for CLAWS Tagger online: (<http://ucrel.lancs.ac.uk/claws7tags.html>)). For the aim of Pronoun constituents analysis, the tags computed are the 3rd person sing. Neuter personal pronoun (it) (PPH1), 3rd person sing. Subjective personal pronoun (he, she) (PPHS1), 1st person sing. Subjective personal pronoun (I) (PPIS1), 1st person plural subjective personal pronoun (we) (PPIS2), 2nd person personal pronoun (you) (PPY), 3rd person plural subjective personal pronoun (they) (PPHS2) forms. The PRONOUNS encoded for CLAWS has seven various tags forming each linguistic constituent of the PRONOUNS;

3rd person sing. Neuter personal pronoun (it) (PPH1), 3rd person sing. Subjective personal pronoun (he, she) (PPHS1), 1st person sing. Subjective personal pronoun (I) (PPIS1), 1st person plural subjective personal pronoun (we) (PPIS2), 2nd person personal pronoun (you) (PPY), 3rd person plural subjective personal pronoun (they) (PPHS2). The recurring linguistics constituents frequency of the PRONOUNS is stated in (Table 2). With respect to the software, it does not do calculation to the total number of PRONOUNS forms employed, it is individually calculated by hand.

Table (2): The number of tagged words

CODES		CLAWS: NUMBER OF TAGGED WORDS
<b>ESSAY WRITING</b>	3SEF1	299 words tagged
	3SEF2	280 words tagged
	3SEF3	814 words tagged
	3SEM1	318 words tagged
	3SEM2	254 words tagged
	3SEM2	320 words tagged
	<b>TOTAL</b>	<b>2285 words tagged</b>

The number of tagged words in female respondent number one were 299 words, the number of tagged words in female respondent number two were 280 words and the number of tagged words in female respondent number three were 814 words. Otherwise in male respondent it is clearly seen the number of tagged words in male respondent number one were 318 words, the number of tagged words in male respondent number two were 254 words and the number of tagged words in male respondent number three were 320 words to give a total number of 2285 tagged words.

Table.3: The use of Personal Pronouns in Essays

CODES		(it) (PPH1)	(he, she) (PPHS1)	(I) (PPIS1)	(we) (PPIS2)	(you) (PPY)	(they) (PPHS2)
<b>ESSAY WRITING</b>	3SEF 1	3	1	6	11	2	0
	3SEF 2	1	0	17	17	1	5
	3SEF 3	10	3	7	59	1	1
	3SEM 1	4	3	4	16	1	0
	3SEM 2	4	0	1	20	0	0
	3SEM 3	1	0	12	15	0	1
	<b>TOTAL</b>	<b>23</b>	<b>7</b>	<b>47</b>	<b>138</b>	<b>5</b>	<b>7</b>

Table.3 shows the usage of Personal Pronouns that is in the six Essays, the table shows the total number of the pronouns used in students’ writings showing the over use of the (PPIS2) pronoun “we” by a total number of 138 mostly used in 3SEF 3 female respondent female number 3. Using Notepad++ which is a free (as in "free speech" and also as in "free beer") source code editor and Notepad replacement that supports several languages. That give a clear picture of the focused pronoun “we” in the learners Essays as follows:

**Essay.1. (3SEF1)**

Find result - 11 hits

Search "we\_PPIS2" (11 hits in 1 file)

F:\PHD\ant conc\CLAWS rxt.tagger\3SEF1.txt (11 hits)

Line 13: town\_NN1 which\_DDQ we\_PPIS2 love\_VV0 most\_RRT . .

Line 16: First\_MD , , we\_PPIS2 went\_VVD to\_II the\_AT beach\_NN1 where\_CS we\_PPIS2

Line 16: First\_MD , , we\_PPIS2 went\_VVD to\_II the\_AT beach\_NN1 where\_CS we\_PPIS2

Line 18: Next\_MD after\_CS we\_PPIS2 had\_VHD finished\_VVN our\_APPGE lunch\_NN1 , ,

Line 19: we\_PPIS2 went\_VVD to\_II the\_AT Cinema\_NN1 to\_TO watch\_VVI a\_AT1 funny\_JJ

Line 23: Then\_RT we\_PPIS2 went\_VVD to\_II the\_AT public\_JJ park\_NN1 where\_CS we\_PPIS2

Line 23: Then\_RT we\_PPIS2 went\_VVD to\_II the\_AT public\_JJ park\_NN1 where\_CS we\_PPIS2

Line 27: When\_CS we\_PPIS2 finished\_VVD we\_PPIS2 went\_VVD to\_TO stay\_VVI the\_AT

Line 27: When\_CS we\_PPIS2 finished\_VVD we\_PPIS2 went\_VVD to\_TO stay\_VVI the\_AT

Line 28: night\_NNT1 at\_II my\_APPGE grandfather\_NN1 home\_RL we\_PPIS2 had\_VHD the\_AT

Line 36: Today\_RT we\_PPIS2 came\_VVD back\_RP to\_II our\_APPGE house\_NN1 still\_RR

## Essay.2. (3SEF2)

Find result - 17 hits

Search "PPIS2" (17 hits in 1 file)

F:\PHD\ant conc\CLAWS rxt.tagger\3SEF2.txt (17 hits)

Line 11: We\_PPIS2 went\_VVD to\_II pyramids\_NN2 . .

Line 12: We\_PPIS2 went\_VVD by\_II my\_APPGE fathers\_NN2 car\_NN1 . .

Line 13: We\_PPIS2 wake\_VV0 up\_RP early\_RR to\_II ready\_JJ and\_CC we\_PPIS2 ate\_VVD the\_AT

Line 13: We\_PPIS2 wake\_VV0 up\_RP early\_RR to\_II ready\_JJ and\_CC we\_PPIS2 ate\_VVD the\_AT

Line 14: breakfast\_NN1 then\_RT we\_PPIS2 moved\_VVD at\_II 6.00\_MC am\_RA and\_CC we\_PPIS2

Line 14: breakfast\_NN1 then\_RT we\_PPIS2 moved\_VVD at\_II 6.00\_MC am\_RA and\_CC we\_PPIS2

Line 15: arrived\_VVD at\_II 10.00am.At\_FO the\_AT first\_MD we\_PPIS2 went\_VVD to\_II the\_AT

Line 24: After\_CS that\_CST we\_PPIS2 went\_VVD with\_IW our\_APPGE parents\_NN2 to\_TO

Line 27: very\_RG good\_JJ . . then\_RT we\_PPIS2 went\_VVD to\_TO see\_VVI a\_AT1 film\_NN1

Line 28: in\_II the\_AT cinema.at\_NNU the\_AT end\_NN1 of\_IO the\_AT day\_NNT1 we\_PPIS2

Line 30: The\_AT next\_MD day\_NNT1 we\_PPIS2 woke\_VVD up\_RP early\_RR to\_TO see\_VVI the\_AT

Line 37: At\_II the\_AT afternoon\_NNT1 we\_PPIS2 went\_VVD back\_RP to\_II the\_AT hotel\_NN1

Line 39: Then\_RT we\_PPIS2 returned\_VVD to\_II home\_NN1 . .

Line 40: Really\_RR we\_PPIS2 all\_DB felt\_VVD a\_AT1 great\_JJ happy\_JJ and\_CC when\_CS

Line 41: we\_PPIS2 saw\_VVD the\_AT pictures\_NN2 we\_PPIS2 feel\_VV0 some\_DD happiness\_NN1

Line 41: we\_PPIS2 saw\_VVD the\_AT pictures\_NN2 we\_PPIS2 feel\_VV0 some\_DD happiness\_NN1

Line 43: We\_PPIS2 thanked\_VVD our\_APPGE parents\_NN2 for\_IF that\_DD1 trip\_NN1 . .

## Essay 3. (3SEF3)

Line 70: our\_APPGE road\_NN1 trip\_NN1 , , we\_PPIS2 saw\_VVD some\_DD old\_JJ villages\_NN2  
 Line 73: As\_CS31 soon\_CS32 as\_CS33 we\_PPIS2 reached\_VVD Cairo\_NP1 . .  
 Line 74: We\_PPIS2 started\_VVD our\_APPGE trip\_NN1 Programme\_NN1 . .  
 Line 78: We\_PPIS2 saw\_VVD some\_DD mummies\_NN2 that\_CST proved\_VVD the\_AT greatness\_NN1  
 Line 80: We\_PPIS2 went\_VVD to\_II opera\_NN1 we\_PPIS2 do\_VD0 nt\_XX understand\_VVI  
 Line 80: We\_PPIS2 went\_VVD to\_II opera\_NN1 we\_PPIS2 do\_VD0 nt\_XX understand\_VVI  
 Line 85: We\_PPIS2 spent\_VVD night\_NNT1 at\_II my\_APPGE relatives\_NN2 flat\_NN1 in\_II  
 Line 86: Cairo\_NP1 we\_PPIS2 left\_VVD back\_RP to\_II our\_APPGE city\_NN1 in\_II the\_AT  
 Line 94: In\_II the\_AT morning\_NNT1 , , we\_PPIS2 went\_VVD to\_II the\_AT sea\_NN1 . .  
 Line 95: We\_PPIS2 swam\_VVD in\_II the\_AT sea\_NN1 and\_CC played\_VVN with\_IW the\_AT  
 Line 97: We\_PPIS2 built\_VVD house\_NN1 on\_II the\_AT sand\_NN1 and\_CC imagined\_VVD  
 Line 99: We\_PPIS2 took\_VVD a\_AT1 boat\_NN1 and\_CC went\_VVD on\_II the\_AT sea\_NN1 . .  
 Line 100: We\_PPIS2 caught\_VVD a\_AT1 lot\_NN1 of\_IO fish\_NN and\_CC enjoyed\_VVD the\_AT  
 Line 102: After\_II midday\_NNT1 we\_PPIS2 went\_VVD to\_II some\_DD historic\_JJ places\_NN2  
 Line 104: We\_PPIS2 went\_VVD to\_II the\_AT castle\_NN1 of\_IO kaytbay\_NN1 and\_CC saw\_VVD  
 Line 105: the\_AT wonderful\_JJ architecture\_NN1 in\_II it\_PPH1 we\_PPIS2 went\_VVD to\_II  
 Line 108: We\_PPIS2 enjoyed\_VVD the\_AT beautiful\_JJ scene\_NN1 of\_IO the\_AT sea\_NN1 . .  
 Line 109: We\_PPIS2 went\_VVD to\_II Alex\_NP1 library\_NN1 , , we\_PPIS2 saw\_VVD many\_DA2  
 Line 109: We\_PPIS2 went\_VVD to\_II Alex\_NP1 library\_NN1 , , we\_PPIS2 saw\_VVD many\_DA2  
 Line 112: At\_II night\_NNT1 we\_PPIS2 went\_VVD to\_II the\_AT funfairs\_NN2 . .  
 Line 113: We\_PPIS2 enjoyed\_VVD many\_DA2 games\_NN2 . .  
 Line 114: We\_PPIS2 bought\_VVD some\_DD toys\_NN2 and\_CC ate\_VVD ice-cream\_NN1 and\_CC  
 Line 116: In\_II the\_AT end\_NN1 we\_PPIS2 went\_VVD home\_RL very\_RG tired\_JJ after\_II  
 Line 117: this\_DD1 long\_JJ day\_NNT1 but\_CCB very\_RG pleased\_JJ as\_CSA we\_PPIS2 found\_VVD  
 Line 118: pleasure\_NN1 in\_II all\_DB things\_NN2 we\_PPIS2 did\_VDD and\_CC all\_DB places\_NN2  
 Line 119: we\_PPIS2 went\_VVD to\_II cinema\_NN1 and\_CC saw\_VVD action\_NN1 film\_NN1 . .  
 Line 120: After\_CS that\_CST we\_PPIS2 went\_VVD to\_II restaurant\_NN1 to\_TO have\_VHI  
 Line 124: life\_NN1 . . in\_II the\_AT end\_NN1 , , we\_PPIS2 can\_VM come\_VVI back\_RP to\_II  
 Line 125: our\_APPGE house\_NN1 and\_CC we\_PPIS2 were\_VBDR very\_RG happy\_JJ to\_TO spend\_VVI

Find result - 59 hits

Search "PPIS2" (59 hits in 1 file)

F:\PHD\ant conc\CLAWS rxt.tagger\3SEF3.txt (59 hits)

Line 7: We\_PPIS2 have\_VHO some\_DD memories\_NN2 with\_IW our\_APPGE families\_NN2 that\_CST  
 Line 8: we\_PPIS2 ca\_VM n't\_XX forget\_VVI all\_DB our\_APPGE lives\_NN2 . .  
 Line 10: forgotten\_VVN whatever\_DDQV we\_PPIS2 become\_VV0 old\_JJ . .  
 Line 14: One\_MC1 of\_IO these\_DD2 days\_NNT2 was\_VBDZ on\_II Friday\_NPD1 we\_PPIS2  
 Line 19: In\_II the\_AT morning\_NNT1 , , we\_PPIS2 went\_VVD to\_II the\_AT sea\_NN1 . .  
 Line 20: We\_PPIS2 swam\_VVD in\_II the\_AT sea\_NN1 and\_CC played\_VVN with\_IW the\_AT  
 Line 22: We\_PPIS2 built\_VVD house\_NN1 on\_II the\_AT sand\_NN1 and\_CC imagined\_VVD  
 Line 24: We\_PPIS2 took\_VVD a\_AT1 boat\_NN1 and\_CC went\_VVD on\_II the\_AT sea\_NN1 . .  
 Line 25: We\_PPIS2 caught\_VVD a\_AT1 lot\_NN1 of\_IO fish\_NN and\_CC enjoyed\_VVD the\_AT  
 Line 27: After\_II midday\_NNT1 we\_PPIS2 went\_VVD to\_II some\_DD historic\_JJ places\_NN2  
 Line 29: We\_PPIS2 went\_VVD to\_II the\_AT castle\_NN1 of\_IO kaytbay\_NN1 and\_CC saw\_VVD  
 Line 30: the\_AT wonderful\_JJ architecture\_NN1 in\_II it\_PPH1 we\_PPIS2 went\_VVD to\_II  
 Line 33: We\_PPIS2 enjoyed\_VVD the\_AT beautiful\_JJ scene\_NN1 of\_IO the\_AT sea\_NN1 . .  
 Line 34: We\_PPIS2 went\_VVD to\_II Alex\_NP1 library\_NN1 , , we\_PPIS2 saw\_VVD many\_DA2  
 Line 34: We\_PPIS2 went\_VVD to\_II Alex\_NP1 library\_NN1 , , we\_PPIS2 saw\_VVD many\_DA2  
 Line 37: At\_II night\_NNT1 we\_PPIS2 went\_VVD to\_II the\_AT funfairs\_NN2 . .  
 Line 38: We\_PPIS2 enjoyed\_VVD many\_DA2 games\_NN2 . .  
 Line 39: We\_PPIS2 bought\_VVD some\_DD toys\_NN2 and\_CC ate\_VVD ice-cream\_NN1 and\_CC  
 Line 41: In\_II the\_AT end\_NN1 we\_PPIS2 went\_VVD home\_RL very\_RG tired\_JJ after\_II  
 Line 42: this\_DD1 long\_JJ day\_NNT1 but\_CCB very\_RG pleased\_JJ as\_CSA we\_PPIS2 found\_VVD  
 Line 43: pleasure\_NN1 in\_II all\_DB things\_NN2 we\_PPIS2 did\_VDD and\_CC all\_DB places\_NN2  
 Line 44: we\_PPIS2 went\_VVD to\_II . .  
 Line 50: We\_PPIS2 all\_DB love\_VV0 each\_PPX221 other\_PPX222 . .  
 Line 51: At\_II hard\_JJ times\_NNT2 , , we\_PPIS2 stand\_VV0 together\_RL as\_CSA one\_MC1  
 Line 57: We\_PPIS2 had\_VHD spent\_VVN such\_DA an\_AT1 amazing\_JJ holiday\_NN1 last\_MD  
 Line 59: There\_EX were\_VBDR many\_DA2 places\_NN2 on\_II where\_RRQ we\_PPIS2 go\_VV0 on\_II  
 Line 64: We\_PPIS2 made\_VVD the\_AT best\_JJT use\_NN1 of\_IO our\_APPGE time\_NNT1 . .  
 Line 65: We\_PPIS2 all\_DB got\_VVD up\_RP early\_RR and\_CC prepared\_VVN for\_IF our\_APPGE  
 Line 67: We\_PPIS2 travelled\_VVD to\_II Cairo\_NP1 on\_II our\_APPGE mini\_JJ bus\_NN1 . .  
 Line 68: We\_PPIS2 reached\_VVD Cairo\_NP1 earlier\_RRR than\_CSN it\_PPH1 was\_VBDZ

Essay.1 (3SEM1)

Find result - 16 hits

Search "PPIS2" (16 hits in 1 file)

F:\PHD\ant conc\CLAWS rxt.tagger\3SEM1.txt (16 hits)

Line 9: We PPIS2 all DB love VV0 each PPX221 other PPX222 . . .  
 Line 10: At II hard JJ times NNT2 , , we PPIS2 stand VV0 together RL as CSA one MC1  
 Line 16: We PPIS2 had VHD spent VVN such DA an AT1 amazing JJ holiday NN1 last MD  
 Line 18: There EX were VBDR many DA2 places NN2 on II where RRQ we PPIS2 go VV0 on II  
 Line 23: We PPIS2 made VVD the AT best JJT use NN1 of IO our APPGE time NNT1 . . .  
 Line 24: We PPIS2 all DB got VVD up RP early RR and CC prepared VVN for IF our APPGE  
 Line 26: We PPIS2 travelled VVD to II Cairo NP1 on II our APPGE mini JJ bus NN1 . . .  
 Line 27: We PPIS2 reached VVD Cairo NP1 earlier RRR than CSN it PPH1 was VBDZ  
 Line 29: our APPGE road NN1 trip NN1 , , we PPIS2 saw VVD some DD old JJ villages NN2  
 Line 32: As CS31 soon CS32 as CS33 we PPIS2 reached VVD Cairo NP1 . . .  
 Line 33: We PPIS2 started VVD our APPGE trip NN1 Programme NN1 . . .  
 Line 37: We PPIS2 saw VVD some DD mummies NN2 that CST proved VVD the AT greatness NN1  
 Line 39: We PPIS2 went VVD to II opera NN1 we PPIS2 do VDO nt XX understand VVI  
 Line 39: We PPIS2 went VVD to II opera NN1 we PPIS2 do VDO nt XX understand VVI  
 Line 44: We PPIS2 spent VVD night NNT1 at II my APPGE relatives NN2 flat NN1 in II  
 Line 45: Cairo NP1 we PPIS2 left VVD back RP to II our APPGE city NN1 in II the AT

## Essay2. (3SEM2)

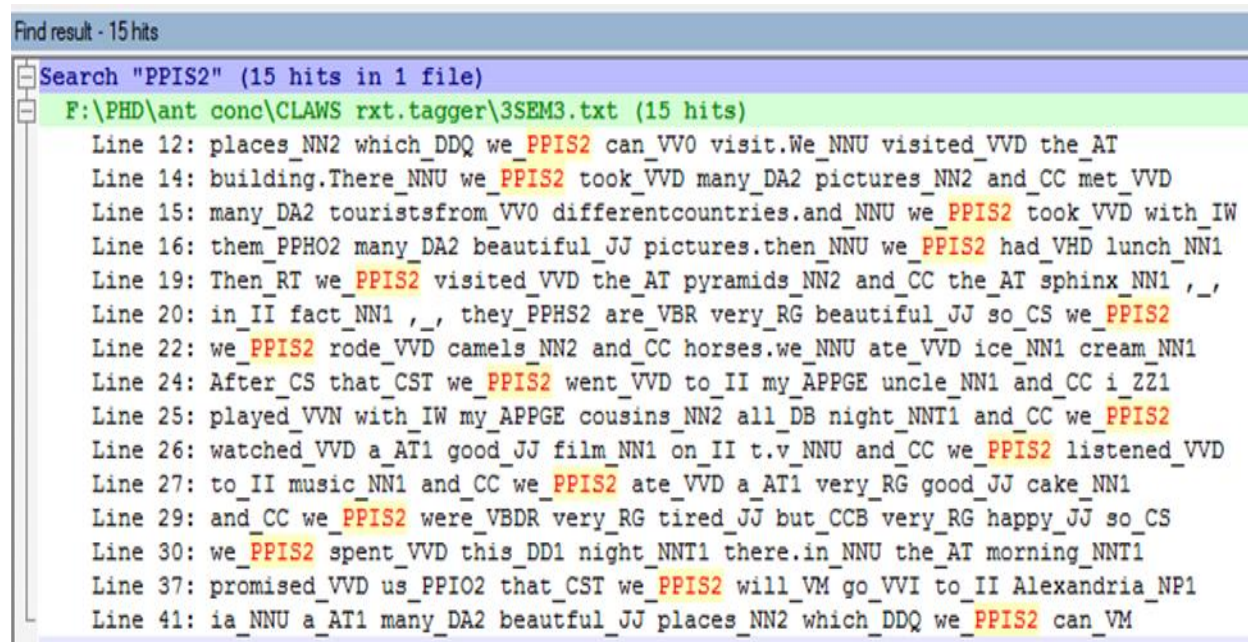
Find result - 20 hits

Search "PPIS2" (20 hits in 1 file)

F:\PHD\ant conc\CLAWS rxt.tagger\3SEM2.txt (20 hits)

Line 11: One MC1 day NNT1 we PPIS2 spent VVD very RG lovely JJ day NNT1 and CC went VVD  
 Line 13: At II the AT begging NN1 of IO the AT day NNT1 , , we PPIS2 woke VVD up RP  
 Line 15: We PPIS2 had VHD breakfast NN1 together RL we PPIS2 have VHO prepared VVN  
 Line 15: We PPIS2 had VHD breakfast NN1 together RL we PPIS2 have VHO prepared VVN  
 Line 16: the AT luggage NN1 and CC tools NN2 . . . we PPIS2 rode VVD the AT car NN1 on II  
 Line 17: our APPGE way NN1 we PPIS2 bought VVD chairs NN2 and CC umbrella NN1 . . .  
 Line 18: We PPIS2 sat VVD on II the AT beach NN1 . . .  
 Line 19: We PPIS2 swam VVD with IW the AT ball NN1 in II the AT sea NN1 . . .  
 Line 20: Then RT we PPIS2 come VV0 back RP into II the AT room NN1 and CC wore VVD  
 Line 22: We PPIS2 went VVD to II and CC had VHD lunch NN1 there RL . . .  
 Line 23: We PPIS2 played VVD tennis NN1 together RL and CC then RT we PPIS2 took VVD  
 Line 23: We PPIS2 played VVD tennis NN1 together RL and CC then RT we PPIS2 took VVD  
 Line 25: After II that DD1 . . . we PPIS2 played VVD tennis NN1 together RL and CC  
 Line 26: then RT we PPIS2 took VVD a AT1 break NN1 in II our APPGE room NN1 . . .  
 Line 27: After CS that CST we PPIS2 went VVD to TO see VVI sights NN2 in II  
 Line 31: In II the AT evening NNT1 , , we PPIS2 went VVD shopping VVG in II the AT  
 Line 33: We PPIS2 went VVD to II cinema NN1 and CC saw VVD action NN1 film NN1 . . .  
 Line 34: After CS that CST we PPIS2 went VVD to II restaurant NN1 to TO have VHI  
 Line 38: life NN1 . . . in II the AT end NN1 , , we PPIS2 can VM come VVI back RP to II  
 Line 39: our APPGE house NN1 and CC we PPIS2 were VBDR very RG happy JJ to TO spend VVI

## Essay3. (3SEM3)



## 5. DISCUSSION

The questions of the present study are answered based on the concordance and the frequency investigation of SSELc.

(1) What are the different forms of Pronouns employed in the collected SSELc corpus?

Practically, it is observed that in the written output of the learners, they utilize a variety of Pronouns forms. Counting on the category of the written assignment, students employ the 3rd person sing. Neuter personal pronoun (it) (PPH1), 3rd person sing. Subjective personal pronoun (he, she) (PPHS1), 1st person sing. Subjective personal pronoun (I) (PPIS1), 1st person plural subjective personal pronoun (we) (PPIS2), 2nd person personal pronoun (you) (PPY), 3rd person plural subjective personal pronoun (they) (PPHS2) forms. The PRONOUNS are illustrated in (Table 4).

Table. (4): The frequency of Personal Pronouns employed in students Essays.

	CODES	(it) (PPH1)	(he, she) (PPHS1)	(I) (PPIS1)	(we) (PPIS2)	(you) (PPY)	(they) (PPHS2)
ESSAY	3SEF1	3	1	6	11	2	0
WRITING	3SEF2	1	0	17	17	1	5
	3SEF3	10	3	7	59	1	1
	3SEM1	4	3	4	16	1	0
	3SEM2	4	0	1	20	0	0
	3SEM2	1	0	12	15	0	1
	<b>TOTAL</b>		<b>23</b>	<b>7</b>	<b>47</b>	<b>138</b>	<b>5</b>

Pedagogically, the students appear to have the ability of employing a variety of choosing words rendering on the various forms of PRONOUNS. The selection of words is reflective of the learners' knowledge of vocabulary and the aptitude to specify the number element of the objects and the subjects of the sentences. Choosing words also showed the written category used, like the choices of words for the essay writing is simple. The investigation is significant to show the pronouns in an individual form in learner corpus because this helps a researcher understand



the structural patterns of a certain category and also support important observations on structural-based linguistic examinations like morphological and syntactic such as syntactic analysis.

(2) In the corpus under investigation, what are the patterns distribution of the forms of the PRONOUN in the SSELc corpus?

The frequency investigation showed that the most frequently employed form of the PRONOUN in different writing categories is 1st person plural subjective personal pronoun (we) (PPIS2).

Table (5): The Concordance hits and the percentage of the PRONOUN forms in the present corpus.

Concordance Hits	140	Total Plots	6
HIT FILE: 1	FILE: 3SEF1.txt		
		No. of Hits = 11	File Length (in chars) = 2964
HIT FILE: 2	FILE: 3SEF2.txt		
		No. of Hits = 17	File Length (in chars) = 2712
HIT FILE: 3	FILE: 3SEF3.txt		
		No. of Hits = 59	File Length (in chars) = 7712
HIT FILE: 4	FILE: 3SEM1.txt		
		No. of Hits = 16	File Length (in chars) = 3100
HIT FILE: 5	FILE: 3SEM2.txt		
		No. of Hits = 20	File Length (in chars) = 2481
HIT FILE: 6	FILE: 3SEM3.txt		
		No. of Hits = 17	File Length (in chars) = 3111

As indicated in Table (5) which shows the total No. of cluster tokens and the percentage of the PRONOUN forms in the current corpus. The ratio was calculated as the percentage of the token in the total number of the PRONOUN forms “we” in the corpus under investigation. The number of hits in file (1):3SEF1 = 11, the number of hits in file (2): 3SEF2=17, the number of hits in file (3): 3SEF3=59, the number of hits in file (4): 3SEM1= 16, the number of hits in file (5):3SEM2=20 and the number of hits in file (6):3SEM3=17. In general, the 1<sup>st</sup> person plural personal subjective pronoun is employed more frequently than other pronouns. The comprehensive indication demonstrates that the percentage is very high for the 1<sup>st</sup> person plural personal subjective pronoun.

As the current study focuses on the use of pronouns, the dominant pronoun used in the learners’ writings was the pronoun “we” this refers to the effect of L1 when the learners write in L2 as in Modern Standard Arabic and also Egyptian Arabic, Subject pronouns in Modern standard and Egyptian Arabic (I, you, we, he, she, we, they) replace a noun and function as the sentence subject ([www.arabic.desert-sky.net/g\\_pronouns.html](http://www.arabic.desert-sky.net/g_pronouns.html)). According to Arabic, the subject pronoun is often dropped. A verb conjugation can you who the subject is. Hence, it's not necessary to utilize the subject pronoun in cases like this except if you want to make emphasis. Yet, in verbless sentences like the two mentioned above, you need the subject pronoun as in table (6).

Table (6): Subject pronouns in Modern Standard and Egyptian Arabic

	English	Standard Arabic	Egyptian Arabic
Singular	I	(ana) أنا	
	you (masc.)	(anta) انت	(inta) انت
	you (fem.)	(anti) انت	(inti) انتى
	he	(howwa) هو	
	she	(heyya) هي	

<b>Dual</b>	<b>we</b>	(naHnu) نحن	
	<b>you</b>	(antuma) أنتما	
	<b>they</b>	(humaa) هما	
<b>Plural</b>	<b>we</b>	(naHnu) نحن	(eHna) احنا
	<b>you (masc.)</b>	(antum) أنتم	(intu) انتو
	<b>you (fem.)</b>	(antunna) أنتن	
	<b>they (masc.)</b>	(homa) هم	(homa) هم
	<b>they (fem.)</b>	(hunna) هن	

Note: There is only one second-person pronoun in English, "you," that is utilized whether you are talking to two people, one person, or more. However, in Arabic, as you can see above, there are feminine and masculine versions of "you," beside to singular, dual (Modern standard Arabic), and plural versions: أنت \ أنتِ if you are addressing one person, أنتما if you are addressing two (in Modern Standard Arabic), and أنتم \ أنتن if you are addressing three or more people. Again, we have to put in our consideration that the dual "you" (أنتما) is exactly the same irrespective to gender. In Modern Standard Arabic, there is a dual version of "they" (هما- that is gender-indiscriminate as well) and feminine and masculine versions of the plural "they" (هم and هن).

Moreover, it is observed that the pronouns in Egyptian Arabic are fewer than Modern Standard Arabic, because it does not have dual pronouns; it has only plural pronouns which are employed to refer to two or more people, of any gender. And the colloquial pronoun انتو and هم are gender-neutral. This greatly affects the writings of the learners when they start to write in English.

Table 7: The total No of Cluster Tokens

Total No. of Cluster Types			2	Total No. of Cluster Tokens		140
Rank	Freq	Range	Cluster			
1	138	6	we_ppis			
2	2	1	we_nnu			

Table (7): shows the total No. of Cluster tokens and the percentage of the PRONOUN forms employed in the current corpus under investigation.

Table 8: The total No of Collocates Tokens

Total No. of Collocate Types: 226				Total No. of Collocate Tokens: 1400	
Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate
1	2	1	1	6.05992	manypictures
2	2	1	1	6.05992	camels
3	6	2	4	5.64489	reached
4	3	1	2	5.64489	finished
5	3	2	1	5.64489	breakfast
6	2	0	2	5.05992	woke
7	1	1	0	5.05992	whatever
8	1	1	0	5.05992	way
9	1	0	1	5.05992	watching
10	1	0	1	5.05992	watched
11	1	0	1	5.05992	wake
12	2	0	2	5.05992	visited
13	3	1	2	5.05992	visit
14	1	1	0	5.05992	v
15	1	1	0	5.05992	umbrella
16	2	0	2	5.05992	travelled
17	1	1	0	5.05992	town
18	1	1	0	5.05992	touristsfrom
19	1	1	0	5.05992	tools
20	1	1	0	5.05992	today

Table 8: The total shows total No of Collocates Tokens and the percentage of the PRONOUN forms utilized in the present corpus.

As showed in Table (8), A great variability in the distributional patterns of the various Pronoun forms is obvious. Generally, the 1<sup>st</sup> person plural subjective personal pronoun (we) (PPIS2) is utilized more frequently than the other forms. The indication of the present study shows that the percentage is very high for the 1<sup>st</sup> person plural subjective personal pronoun comparing to the other forms of the pronoun.

The tendency to rely so much on 1<sup>st</sup> person plural subjective personal pronoun (we) (PPIS2) is observed in all the different categories used for the compilation of the corpus under investigation. The 1<sup>st</sup> person plural subjective personal pronoun (we) (PPIS2) is used 138 OUT OF 227 equal to (60.79) times in the essays.

It is noticed that PRONOUN forms tend to be utilized more frequently as sentences subject. Hence, PRONOUN forms are bounded to be employed by the writers to start a sentence. Nevertheless, the distributional examination shows the overreliance on the 1<sup>st</sup> person plural subjective personal pronoun (we) (PPIS2) form in all the categories in the current corpus. A high percentage of the 1<sup>st</sup> person plural subjective personal pronoun (we) (PPIS2) has showed that these learners are acquainted with the singular and the plural form of the PRONOUN constituent. Also, the indication shows that the knowledge of the application of grammar rules related to the use of the PRONOUN constituent.

Moreover, the distributional investigation indicates that essay writing has many of tokens representing 1<sup>st</sup> person plural subjective personal pronoun (we) (PPIS2). The implication is that those learners may have depended on 1<sup>st</sup> person plural subjective personal pronoun (we) (PPIS2) when they write to give a description to what want to say when describing past events or tell a stories about certain actions they done when the travelled with their families' even it is very normal when you are sitting with your family watching a TV program to find someone of the family member say "we" want to change the channel while if you ask her /him are all of you want to change the channel he/she say no I want ... then you will find yourself ask again why do you use the pronoun "we" but this is normal in Modern Standard Arabic or even Egyptian Arabic so when they write they were influenced by their mother tongue.

## 6. CONCLUSION

The investigation of the present study has been utilized to examine a specific linguistics constituent; the PRONOUN. To differentiate the number of each employed PRONOUN forms, its frequency in the annotated corpus is tabulated and calculated. The most used pronoun form is pointed out. Also, the distributional patterns of each noun form in the gathered learner corpus is investigated. Hence, the outcomes are significant for answering the research questions of the present study.

Yet, the present study is concentrated on a small sample of a learner corpus that is gathered particularly for the investigation. The results might not be appropriate to be applied and generalized on another larger corpus. The

corpus sample under analysis is limited categories. Nevertheless, it is significant to compile and analyze a wider range of various genres in different areas such as the context of business and management.

However, the outcomes of this modest study cast some vital insights into the employment of linguistics constituents produced by second language learners Essay context. After the analysis, it is observed that there are reliably various forms of pronoun manipulated by learners so as to develop sentences. The formations of the PRONOUN constituent contribute to the understanding of the deliberated written context texts since these pronoun formations are manipulated as sentences' subjects. The investigation has highlighted the over-dependency on the 1<sup>st</sup> person plural subjective personal pronoun (we) (PPIS2). It would be interesting to further investigate how the 1<sup>st</sup> person plural subjective personal pronoun (we) (PPIS2) contributes the correct employment of the grammar rule associated with the agreement of subject-verb. The investigation also reveals that learners are likely to utilize more pronoun forms in their essay assignments.

## REFERENCES

- [1] Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. ... *Learner Corpora, Second Language Acquisition and ...*
- [2] Chan, S., Bee, W., Wong, B., & Teo, H. (2004). The acquisition of the English inflectional-s Morphemes by Young L1 Chinese Speakers. *GEMA: Online Journal of ...*
- [3] Garside, R. (1987). The CLAWS word-tagging system. *The Computational Analysis of English: A Corpus-Based Approach*, 30–41.
- [4] Garside, R., Fligelstone, S., & Botley, S. (1997). Discourse annotation: anaphoric relations in corpora. ... : *Linguistic Information from Computer Text Corpora*.
- [5] Garside, Roger, & Smith, N. (1997). A hybrid grammatical tagger: CLAWS4. *Corpus Annotation: Linguistic Information from Computer Text Corpora*, 102–121.
- [6] Goldwater, S., & McClosky, D. (2005). Improving statistical MT through morphological analysis. *Proceedings of the Conference on Human*.
- [7] Kaur, M., & Shamsudin, S. (2011). Extracting noun forms: A lesson learnt. *International Journal of Language Studies (IJLS)*, 5(4), 2011–2019.
- [8] Kaur, M., Singh, S., Pendidikan, F., Teknologi, U., Sarimah, M., Fakulti, S., ... Manusia, S. (2014). Arriving at the interface of pragmatic and grammatical competence – a situation analysis of English language learners' needs.
- [9] Lee, Y. (2004). Morphological analysis for statistical machine translation. *Proceedings of HLT-NAACL 2004: Short Papers*. Retrieved from <http://dl.acm.org/citation.cfm?id=1613999>
- [10] Leech, G., Garside, R., & Bryant, M. (1994). CLAWS4: the tagging of the British National Corpus. *Proceedings of the 15th Conference on ...*
- [11] Popović, M., Ney, H., & Gispert, A. De. (2006). Morpho-syntactic information for automatic error analysis of statistical machine translation output. ... *Machine Translation*.
- [12] Popović, M., Vilar, D., & Ney, H. (2005). Augmenting a small parallel text with morpho-syntactic language resources for Serbian-English statistical machine translation.
- [13] Yusuf, Y., & Qismullah Yusuf, Y. (2009). A Corpus-Based Linguistics Analysis on Written Corpus: Colligation of " TO " and " FOR ". *Journal of Language and Linguistic Studies*, 5(2).
- [14] Zhang, Y. (2008). A Corpus-based Study of the Forms and Functions of BE in the Interlanguage Grammars of Chinese Learners of English.