**Global Journal of Arts Humanity and Social Sciences ISSN: 2583-2034**

Glob. J.Arts.Humanit.Soc.Sci
ISSN: 2583-2034
Vol-2 Iss-10, page 731-738

# Linguistics in Big Data

**BY**

**Namkil Kang**
Far East University South Korea

Article History
Received: 30/09/2022
Accepted: 03/10/2022
Published: 05/10/2022

**Corresponding author:**

**Namkil Kang**

**Abstract**
The main purpose of this paper is to analyze 576 KCI (Korea Citation Index) articles from 2002 to 2022. A major point to note is that the word *linguistic* was the most widely used one, followed by the word *language*, the word *study*, the word *Linguistics*, the word *research*, and the word *paper*. A further point to note is that topic 7 was the most preferred one for authors, followed by topic 8 (topic 10), topic 5, and topic 9, in descending order. It is interesting to note that as the 1st keyword, the word *linguistic* was the most preferred one. With respect to degree (the frequency of documents), it is worth pointing out that the word *Linguistics* was the most preferred one for authors, followed by the word *linguistic*, the word *study*, and the word *language*. Talking about the word *Linguistics*, it occurred in 404 articles, which is the highest. This in turn shows that authors preferred using the word *Linguistics* to using the other words. Finally, the visualization of words occurring with the word *linguistics* provides us with the picture of which words are linked to it.

**Keywords:** article, KCI, big data, degree, frequency, NetMiner

## 1. Introduction

The main purpose of this paper is to analyze 576 KCI articles published from 2002 to 2022 that are related with the keyword *linguistics*. First, we classify all of them into time period. Second, we inquire into the frequency of words occurred in 576 KCI articles (titles, abstracts, and keywords). Third, we look into 10 topics in which each topic is constituted by 5 keywords. Articles are constituted by topics, which are formed by words. Fourth, we consider the number of articles in which each topic was used. Fifth, we consider degree (the term of NetMiner), which indicates "In how many papers did a word appear?" Sixth, we capture main words neighboring with the word *linguistics* in terms of the visualization of those words. The organization of this paper is as follows. In section 3.2, we argue that the word *linguistic* was the most widely used by authors, followed by the word *language*, the word *study*, the word *Linguistics*, the word *research*, and the word *paper*, in that order. In section 3.3, we further argue that topic 7 was the most preferred one for authors, followed by topic 8 (topic 10), topic 5, and topic 9, in descending order. We also maintain that the word *linguistic* was the most preferred one as the 1st keyword. In section 3.4, we contend that the word *Linguistics* was the most preferred one for authors, followed by the word *linguistic*, the word *study*, and the word *language*. When it comes to the word *Linguistics*, it appeared in 404 articles, which is the highest. This in turn suggests that authors preferred using the word *Linguistics* to using the other words. In section 3.5, we visualize words neighboring with the word *linguistics*.

## 2. Methods

The main goal of this paper is to analyze 576 KCI articles published from 2002 to 2022. We used the Biblio data collector to collect them. We analyzed all of them in terms of the software package NetMiner. The goal of this paper is to answer the following questions: Can we provide information on 576 KCI articles including their proportion and cumulative proportion? What does the frequency of words occurred in 576 KCI articles indicate? Can we provide topics that are formed by main keywords? Can we provide information on degree (the term of NetMiner)? Can we capture main words neighboring with the word *linguistics* (their visualization)?

## 3. Results

### 3.1. Frequency of 576 articles

In this section, we aim to provide information on 576 KCI articles including their proportion and cumulative proportion. Table 1 shows the number of articles, their proportion, and their cumulative proportion:

**Table 1 Frequency of articles published from 2002 to 2022**

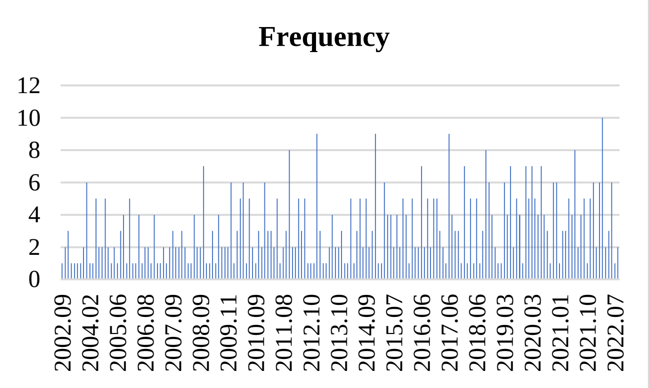
| Value | Frequency | Proportion | Cumulative Proportion |
|---|---|---|---|
| 2002.09 | 1 | 0.002 | 0.002 |
| 2002.12 | 2 | 0.003 | 0.005 |
| 2003.03 | 3 | 0.005 | 0.01 |
| 2003.05 | 1 | 0.002 | 0.012 |
| 2003.06 | 1 | 0.002 | 0.014 |
| 2003.08 | 1 | 0.002 | 0.016 |
| 2003.09 | 1 | 0.002 | 0.017 |
| 2003.11 | 2 | 0.003 | 0.021 |
| 2003.12 | 6 | 0.01 | 0.031 |
| 2004.02 | 1 | 0.002 | 0.033 |
| 2004.05 | 1 | 0.002 | 0.035 |
| 2004.06 | 5 | 0.009 | 0.043 |
| 2004.08 | 2 | 0.003 | 0.047 |
| 2004.09 | 2 | 0.003 | 0.05 |
| 2004.12 | 5 | 0.009 | 0.059 |
| 2005.02 | 2 | 0.003 | 0.063 |
| 2005.03 | 1 | 0.002 | 0.064 |
| 2005.05 | 2 | 0.003 | 0.068 |
| 2005.06 | 1 | 0.002 | 0.069 |
| 2005.08 | 3 | 0.005 | 0.075 |
| 2005.09 | 4 | 0.007 | 0.082 |
| 2005.11 | 1 | 0.002 | 0.083 |
| 2005.12 | 5 | 0.009 | 0.092 |
| 2006.01 | 1 | 0.002 | 0.094 |
| 2006.02 | 1 | 0.002 | 0.095 |
| 2006.06 | 4 | 0.007 | 0.102 |
| 2006.07 | 1 | 0.002 | 0.104 |
| 2006.08 | 2 | 0.003 | 0.108 |
| 2006.09 | 2 | 0.003 | 0.111 |
| 2006.11 | 1 | 0.002 | 0.113 |
| 2006.12 | 4 | 0.007 | 0.12 |
| 2007.03 | 1 | 0.002 | 0.122 |
| 2007.04 | 1 | 0.002 | 0.123 |
| 2007.05 | 2 | 0.003 | 0.127 |
| 2007.06 | 1 | 0.002 | 0.128 |
| 2007.08 | 2 | 0.003 | 0.132 |
| 2007.09 | 3 | 0.005 | 0.137 |
| 2007.10 | 2 | 0.003 | 0.141 |
| 2007.11 | 2 | 0.003 | 0.144 |
| 2007.12 | 3 | 0.005 | 0.149 |
| 2008.02 | 2 | 0.003 | 0.153 |
| 2008.03 | 1 | 0.002 | 0.155 |
| 2008.05 | 1 | 0.002 | 0.156 |
| 2008.06 | 4 | 0.007 | 0.163 |
| 2008.08 | 2 | 0.003 | 0.167 |
| 2008.09 | 2 | 0.003 | 0.17 |
| 2008.12 | 7 | 0.012 | 0.182 |
| 2009.02 | 1 | 0.002 | 0.184 |
| 2009.03 | 1 | 0.002 | 0.186 |
| 2009.04 | 3 | 0.005 | 0.191 |
| 2009.05 | 1 | 0.002 | 0.193 |
| 2009.06 | 4 | 0.007 | 0.2 |
| 2009.08 | 2 | 0.003 | 0.203 |
| 2009.10 | 2 | 0.003 | 0.207 |
| 2009.11 | 2 | 0.003 | 0.21 |
| 2009.12 | 6 | 0.01 | 0.22 |
| 2010.01 | 1 | 0.002 | 0.222 |
| 2010.02 | 3 | 0.005 | 0.227 |
| 2010.03 | 5 | 0.009 | 0.236 |
| 2010.04 | 6 | 0.01 | 0.247 |
| 2010.05 | 1 | 0.002 | 0.248 |
| 2010.06 | 5 | 0.009 | 0.257 |
| 2010.08 | 2 | 0.003 | 0.26 |
| 2010.09 | 1 | 0.002 | 0.262 |
| 2010.10 | 3 | 0.005 | 0.267 |
| 2010.11 | 2 | 0.003 | 0.271 |
| 2010.12 | 6 | 0.01 | 0.281 |
| 2011.01 | 3 | 0.005 | 0.286 |
| 2011.02 | 3 | 0.005 | 0.292 |
| 2011.04 | 2 | 0.003 | 0.295 |
| 2011.05 | 5 | 0.009 | 0.304 |
| 2011.06 | 1 | 0.002 | 0.306 |

| | | | |
|---|---|---|---|
| 2011.08 | 2 | 0.003 | 0.309 |
| 2011.10 | 3 | 0.005 | 0.314 |
| 2011.12 | 8 | 0.014 | 0.328 |
| 2012.02 | 2 | 0.003 | 0.332 |
| 2012.03 | 2 | 0.003 | 0.335 |
| 2012.04 | 5 | 0.009 | 0.344 |
| 2012.06 | 3 | 0.005 | 0.349 |
| 2012.08 | 5 | 0.009 | 0.358 |
| 2012.09 | 1 | 0.002 | 0.359 |
| 2012.10 | 1 | 0.002 | 0.361 |
| 2012.11 | 1 | 0.002 | 0.363 |
| 2012.12 | 9 | 0.016 | 0.378 |
| 2013.02 | 3 | 0.005 | 0.384 |
| 2013.03 | 1 | 0.002 | 0.385 |
| 2013.04 | 1 | 0.002 | 0.387 |
| 2013.06 | 2 | 0.003 | 0.391 |
| 2013.08 | 4 | 0.007 | 0.398 |
| 2013.09 | 2 | 0.003 | 0.401 |
| 2013.10 | 2 | 0.003 | 0.405 |
| 2013.11 | 3 | 0.005 | 0.41 |
| 2013.12 | 1 | 0.002 | 0.411 |
| 2014.02 | 1 | 0.002 | 0.413 |
| 2014.03 | 5 | 0.009 | 0.422 |
| 2014.04 | 1 | 0.002 | 0.424 |
| 2014.05 | 3 | 0.005 | 0.429 |
| 2014.06 | 5 | 0.009 | 0.437 |
| 2014.08 | 2 | 0.003 | 0.441 |
| 2014.09 | 5 | 0.009 | 0.45 |
| 2014.10 | 2 | 0.003 | 0.453 |
| 2014.11 | 3 | 0.005 | 0.458 |
| 2014.12 | 9 | 0.016 | 0.474 |
| 2015.02 | 1 | 0.002 | 0.476 |
| 2015.03 | 1 | 0.002 | 0.477 |
| 2015.04 | 6 | 0.01 | 0.488 |
| 2015.05 | 4 | 0.007 | 0.495 |
| 2015.06 | 4 | 0.007 | 0.502 |
| 2015.07 | 2 | 0.003 | 0.505 |
| 2015.08 | 4 | 0.007 | 0.512 |
| 2015.09 | 2 | 0.003 | 0.516 |
| 2015.11 | 5 | 0.009 | 0.524 |
| 2015.12 | 4 | 0.007 | 0.531 |
| 2016.01 | 1 | 0.002 | 0.533 |
| 2016.03 | 5 | 0.009 | 0.542 |
| 2016.04 | 2 | 0.003 | 0.545 |
| 2016.05 | 2 | 0.003 | 0.549 |
| 2016.06 | 7 | 0.012 | 0.561 |
| 2016.08 | 2 | 0.003 | 0.564 |
| 2016.09 | 5 | 0.009 | 0.573 |
| 2016.10 | 2 | 0.003 | 0.576 |
| 2016.11 | 5 | 0.009 | 0.585 |
| 2016.12 | 5 | 0.009 | 0.594 |
| 2017.02 | 3 | 0.005 | 0.599 |
| 2017.04 | 2 | 0.003 | 0.602 |
| 2017.05 | 1 | 0.002 | 0.604 |
| 2017.06 | 9 | 0.016 | 0.62 |
| 2017.08 | 4 | 0.007 | 0.627 |
| 2017.09 | 3 | 0.005 | 0.632 |
| 2017.10 | 3 | 0.005 | 0.637 |
| 2017.11 | 1 | 0.002 | 0.639 |
| 2017.12 | 7 | 0.012 | 0.651 |
| 2018.02 | 1 | 0.002 | 0.653 |
| 2018.03 | 5 | 0.009 | 0.661 |
| 2018.05 | 1 | 0.002 | 0.663 |
| 2018.06 | 5 | 0.009 | 0.672 |
| 2018.07 | 1 | 0.002 | 0.674 |
| 2018.08 | 3 | 0.005 | 0.679 |
| 2018.09 | 8 | 0.014 | 0.693 |
| 2018.10 | 6 | 0.01 | 0.703 |
| 2018.11 | 4 | 0.007 | 0.71 |
| 2018.12 | 2 | 0.003 | 0.714 |
| 2019.01 | 1 | 0.002 | 0.715 |
| 2019.02 | 1 | 0.002 | 0.717 |
| 2019.03 | 6 | 0.01 | 0.727 |
| 2019.05 | 4 | 0.007 | 0.734 |
| 2019.06 | 7 | 0.012 | 0.747 |
| 2019.07 | 2 | 0.003 | 0.75 |

| | | | |
|---|---|---|---|
| 2019.08 | 5 | 0.009 | 0.759 |
| 2019.09 | 4 | 0.007 | 0.766 |
| 2019.11 | 1 | 0.002 | 0.767 |
| 2019.12 | 7 | 0.012 | 0.78 |
| 2020.02 | 5 | 0.009 | 0.788 |
| 2020.03 | 7 | 0.012 | 0.8 |
| 2020.04 | 5 | 0.009 | 0.809 |
| 2020.05 | 4 | 0.007 | 0.816 |
| 2020.06 | 7 | 0.012 | 0.828 |
| 2020.08 | 4 | 0.007 | 0.835 |
| 2020.09 | 3 | 0.005 | 0.84 |
| 2020.10 | 1 | 0.002 | 0.842 |
| 2020.11 | 6 | 0.01 | 0.852 |
| 2020.12 | 6 | 0.01 | 0.863 |
| 2021.01 | 1 | 0.002 | 0.865 |
| 2021.02 | 3 | 0.005 | 0.87 |
| 2021.03 | 3 | 0.005 | 0.875 |
| 2021.04 | 5 | 0.009 | 0.884 |
| 2021.05 | 4 | 0.007 | 0.891 |
| 2021.06 | 8 | 0.014 | 0.905 |
| 2021.07 | 2 | 0.003 | 0.908 |
| 2021.08 | 4 | 0.007 | 0.915 |
| 2021.09 | 5 | 0.009 | 0.924 |
| 2021.10 | 1 | 0.002 | 0.925 |
| 2021.11 | 5 | 0.009 | 0.934 |
| 2021.12 | 6 | 0.01 | 0.944 |
| 2022.01 | 2 | 0.003 | 0.948 |
| 2022.02 | 6 | 0.01 | 0.958 |
| 2022.03 | 10 | 0.017 | 0.976 |
| 2022.04 | 2 | 0.003 | 0.979 |
| 2022.05 | 3 | 0.005 | 0.984 |
| 2022.06 | 6 | 0.01 | 0.995 |
| 2022.07 | 1 | 0.002 | 0.997 |
| 2022.08 | 2 | 0.003 | 1 |
| Total | 576 | 1 | |

It is important to note that 10 KCI papers were published in March in 2020, which rank first (the highest). Their proportion and cumulative proportion are 0.01 and 0.958, respectively. As illustrated in Table 1, in December in 2014, 9 KCI papers were published, which rank second. Their proportion and cumulative proportion are 0.016 and 0.474, respectively. Likewise, in December in 2012, 9 KCI papers were published (the second highest). Their proportion is 0.016 and their cumulative proportion is 0.378. Also, in June in 2017, the same number was published, which also ranks second. It is worth observing, on the other hand, that 8 KCI papers in 2011, 2018, and 2021 were published, which rank third (the third highest). The proportion and cumulative proportion of the article that was published in December in 2011 were 0.014 and 0.328, respectively. On the other hand, talking about the proportion and cumulative proportion of the article published in September in 2018, they are 0.014 and 0.693, respectively. 8 KCI articles were published in August in 2021 whose proportion and cumulative proportion are 0.014 and 0.905, respectively. It is worthwhile noting that 7 KCI papers were published in 2008, 2016, 2019, and 2020, which rank fourth (the fourth highest). It is interesting to note, on the other hand, that 6 KCI papers were also published in 2003, 2009, 2010, 2018, 2019, 2020, 2021, and 2022 that are the fifth highest in number. Now the following graph briefly illustrates the frequency of 576 articles published from 2002 to 2022:

**Figure 1 Frequency of 576 articles published from 2002 to 2022**



### 3.2. A Frequency Analysis of Words

In this section, we look into the frequency of words used in 576 KCI articles (titles, abstracts, and keywords). The list was cut off in the top 50. Table 2 shows the frequency of words that occur in 576 KCI articles:

**Table 2 Frequency of words**

| Number | Word | Tag | Frequency |
|---|---|---|---|
| **1** | linguistic | Adjective | 1,106 |
| **2** | language | Noun | 1,091 |
| **3** | study | Noun | 835 |
| **4** | Linguistics | Noun | 651 |
| **5** | research | Noun | 565 |
| **6** | paper | Noun | 356 |
| **7** | Korean | Noun | 348 |
| **8** | text | Noun | 323 |
| **9** | analysis | Noun | 288 |
| **10** | word | Noun | 286 |

| | | | |
|---|---|---|---|
| **11** | theory | Noun | 253 |
| **12** | education | Noun | 238 |
| **13** | meaning | Noun | 232 |
| **14** | English | Noun | 223 |
| **15** | grammar | Noun | 211 |
| **16** | article | Noun | 181 |
| **17** | result | Noun | 172 |
| **18** | field | Noun | 168 |
| **19** | method | Noun | 152 |
| **20** | perspective | Noun | 151 |
| **21** | structure | Noun | 145 |
| **22** | translation | Noun | 145 |
| **23** | metaphor | Noun | 139 |
| **24** | system | Noun | 139 |
| **25** | approach | Noun | 138 |
| **26** | purpose | Noun | 134 |
| **27** | concept | Noun | 132 |
| **28** | term | Noun | 131 |
| **29** | Korea | Noun | 128 |
| **30** | time | Noun | 128 |
| **31** | student | Noun | 122 |
| **32** | process | Noun | 120 |
| **33** | type | Noun | 118 |
| **34** | characteristic | Noun | 113 |
| **35** | literature | Noun | 111 |
| **36** | use | Noun | 110 |
| **37** | datum | Noun | 107 |
| **38** | function | Noun | 106 |
| **39** | verb | Noun | 106 |
| **40** | Study | Noun | 103 |
| **41** | history | Noun | 102 |
| **42** | Language | Noun | 99 |
| **43** | knowledge | Noun | 96 |
| **44** | area | Noun | 94 |
| **45** | way | Noun | 94 |
| **46** | content | Noun | 93 |
| **47** | writing | Noun | 93 |
| **48** | difference | Noun | 92 |
| **49** | noun | Noun | 90 |
| **50** | trend | Noun | 90 |

It is significant to note that the word *linguistic* has the highest frequency and the highest proportion. More specifically, the frequency of the word *linguistic* is 1,106 tokens. This in turn implies that the word *linguistic* was the most preferred one for authors for 20 years from 2002 to 2022. It is worth mentioning, on the other hand, that the word *language* was the second most widely used one (1,091 tokens). Quite interestingly, the word *study* was the third most frequently used one (835 tokens). Note that the frequency of the word *Linguistics* is 651 tokens, which rank fourth (the fourth highest). It should be pointed out that the frequency of the word *research* is 565 tokens, which are the fifth highest. Simply put, the word *research* was the fifth most preferred one for authors. It is worthwhile pointing out that the word *paper* ranks sixth (356 tokens). From all of this, it is evident that the word *linguistic* was the most widely used by authors, followed by the word *language*, the word *study*, the word *Linguistics*, the word *research*, and the word *paper*, in that order. It must be noted, however, that the word *text* ranks eighth (323 tokens). More interestingly, the word *analysis* was the ninth most preferred one for authors (288 tokens). That is to say, it is the ninth highest among words occurred in 576 KCI papers. With respect to the word *grammar*, it is worthwhile pointing out that it was the fifteenth widely used one (211 tokens). Additionally, it should be mentioned that the word *method* is the sixteenth frequently used one (152 tokens). Finally, it must be pointed out that the frequency of the word *structure* is 145 tokens, which rank seventeenth. This in turn implies that the word *structure* was the seventh most widely used one (145 tokens). We thus conclude that the word *linguistic* was the most preferred one for authors.

### 3.3. Topic Information

In this section, we inquire into 10 topics in which each topic is constituted by 5 keywords. Table 3 shows each topic which is formed by 5 keywords:

**Table 3 Topic Information**

| | **1st Keyword** | **2nd Keyword** | **3rd Keyword** | **4th Keyword** | **5th Keyword** |
|---|---|---|---|---|---|
| **Topic-1** | text | type | structure | process | linguistic |
| **Topic-2** | education | grammar | linguistic | student | content |
| **Topic-3** | analysis | study | text | research | datum |
| **Topic-4** | word | Linguistics | system | part | structure |
| **Topic-5** | linguistic | study | research | theory | language |
| **Topic-6** | English | study | research | linguistic | writing |
| **Topic-7** | linguistic | meaning | language | Linguistics | study |
| **Topic-8** | language | linguistic | research | study | Linguistics |

| | | | | | |
|---|---|---|---|---|---|
| **Topic-9** | study | language | Korean | Linguistics | word |
| **Topic-10** | research | linguistic | study | Linguistics | paper |

It is interesting to point out that 5 keywords such as *text*, *type*, *structure*, *process*, and *linguistic* constitute topic 1. It must be noted that in topic 1, the word *text* appears as the 1st keyword, whereas it occurs as the 3rd keyword in topic 3. This in turn suggests that this keyword as the 1st topic was not much used. It is worth mentioning, on the other hand, that 5 keywords such as *education*, *grammar*, *linguistic*, *student*, and *content* form topic 2. Quite interestingly, in topic 2, the word *linguistic* appears as the 3rd keyword. It is worthwhile noting that topic 5 is formed by 5 keywords such as *linguistic*, *study*, *research*, *theory*, and *language*. On the other hand, 5 keywords such as *research*, *linguistic*, *study*, *Linguistics*, and *paper* constitute topic 10. It is important to note that as the 1st keyword, the word *linguistic* ranks first (the highest), as indicated in Table 3. This in turn indicates that authors preferred using the word *linguistic* rather than using the other words. It is worthwhile pointing out that as the 2nd keyword, the word *study* was the second most frequently used one. This in turn suggests that as the 2nd keyword, the word *study* was the second most preferred one for authors. When it comes to the 3rd keyword, the word *research* ranks first (the highest). Talking about the 4th keyword, the word *Linguistics* were the most preferred one for author. Finally, it should be mentioned that in 10 topics, the word *linguistic* was the most frequently used one.

Now, attention is paid to topics and the frequency of documents:

**Table 4 Topics and the frequency of documents**

| | # of documents |
|---|---|
| **Topic-1** | 26 |
| **Topic-2** | 24 |
| **Topic-3** | 29 |
| **Topic-4** | 50 |
| **Topic-5** | 79 |
| **Topic-6** | 40 |
| **Topic-7** | 91 |
| **Topic-8** | 84 |
| **Topic-9** | 69 |
| **Topic-10** | 84 |

It is significant to note that as illustrated in Table 4, topic 7 that is constituted by 5 keywords such as *linguistic*, *meaning*, *language*, *Linguistics*, and *study* appeared in 91 articles. Table 4 clearly indicates that this figure is the highest. From this, it is clear that authors preferred using the keywords *linguistic*, *meaning*, *language*, *Linguistics*, and *study* to using the other keywords. It is worth observing that topic 8 occurred in 84 articles. As observed earlier, the keywords *language*, *linguistic*, *research*, *study*, and *Linguistics* constitute topic 8. Exactly the same can be said of topic 10. Topic 10 also occurred in 84 articles. However, topic 10 is different from topic 8 in that the former is constituted by five keywords such as *research*, *linguistic*, *study*, *Linguistics*, and *paper*. It is worth pointing out that topic 5 occurred in 79 articles, which rank third. 5 keywords such as *linguistic*, *study*, *research*, *theory*, and *language* consist of topic 5. Also, it should be pointed out that topic 9 that is formed by *study*, *language*, *Korean*, *Linguistics*, and *word* occurred in 69 articles. From all of this, it is clear that topic 7 was the most preferred one for authors, followed by topic 8 (topic 10), topic 5, and topic 9, in descending order. Finally, it must be noted that as indicated in Table 4, topic 2 occurred in 24 articles, which ranks tenth. This in turn indicates that topic 2 was the least preferred one. As observed earlier, this topic is constituted by five keywords such as *education*, *grammar*, *linguistics*, *student*, and *content*.

### 3.4. Degree

In what follows, we are concerned with degree (the term of NetMiner). This indicates “In how many documents did a particular word occur?” Table 5 shows degree, namely the frequency of articles:

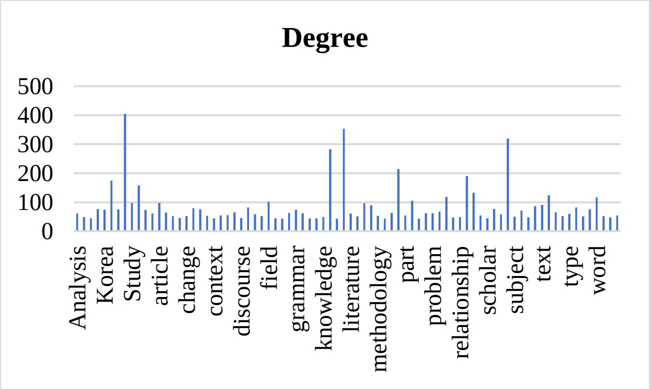
**Table 5 Degree**

| Number | Word | Degree |
|---|---|---|
| **1** | Linguistics | 404 |
| **2** | linguistic | 354 |
| **3** | study | 319 |
| **4** | language | 283 |
| **5** | paper | 215 |
| **6** | research | 190 |
| **7** | Korean | 175 |
| **8** | analysis | 158 |
| **9** | result | 133 |
| **10** | theory | 124 |
| **11** | purpose | 118 |
| **12** | word | 117 |
| **13** | perspective | 105 |
| **14** | field | 102 |
| **15** | article | 98 |
| **16** | Study | 98 |
| **17** | meaning | 96 |
| **18** | text | 91 |
| **19** | method | 90 |
| **20** | term | 86 |
| **21** | education | 82 |
| **22** | use | 82 |
| **23** | characteristic | 80 |
| **24** | English | 77 |

| | | |
|---|---|---|
| **25** | structure | 77 |
| **26** | concept | 76 |
| **27** | Language | 76 |
| **28** | way | 76 |
| **29** | grammar | 74 |
| **30** | Korea | 74 |
| **31** | approach | 73 |
| **32** | system | 71 |
| **33** | process | 68 |
| **34** | time | 66 |
| **35** | difference | 65 |
| **36** | aspect | 64 |
| **37** | function | 63 |
| **38** | order | 63 |
| **39** | history | 62 |
| **40** | point | 62 |
| **41** | problem | 62 |
| **41** | Analysis | 61 |
| **43** | area | 61 |
| **44** | literature | 61 |
| **45** | type | 60 |
| **46** | example | 59 |
| **47** | student | 59 |
| **48** | development | 56 |
| **49** | datum | 55 |
| **50** | year | 55 |
| **51** | part | 54 |
| **52** | role | 54 |
| **53** | content | 53 |
| **54** | methodology | 53 |
| **55** | basis | 52 |
| **56** | change | 52 |
| **57** | expression | 52 |
| **58** | trend | 52 |
| **59** | work | 52 |
| **60** | material | 51 |
| **61** | view | 51 |
| **62** | subject | 50 |
| **63** | Chinese | 49 |
| **64** | knowledge | 49 |
| **65** | relationship | 49 |
| **66** | relation | 48 |
| **67** | teaching | 48 |
| **68** | writing | 48 |
| **69** | case | 46 |
| **70** | discourse | 46 |
| **71** | Cognitive | 45 |
| **72** | context | 45 |
| **73** | finding | 45 |
| **74** | information | 45 |
| **75** | issue | 45 |
| **76** | scholar | 45 |
| **77** | form | 44 |
| **78** | phenomenon | 44 |
| **79** | level | 43 |
| **80** | number | 43 |
| **81** | addition | 41 |
| **82** | model | 41 |
| **83** | understanding | 41 |
| **84** | direction | 40 |
| **85** | discussion | 40 |
| **86** | need | 40 |
| **87** | researcher | 40 |
| **88** | science | 40 |
| **89** | sentence | 40 |
| **90** | construction | 39 |
| **91** | feature | 39 |
| **92** | period | 39 |
| **93** | learner | 38 |
| **94** | question | 38 |
| **95** | society | 38 |
| **96** | verb | 38 |
| **97** | corpus | 37 |
| **98** | implication | 37 |
| **99** | topic | 37 |

It is significant to note that the word *Linguistics* has the highest frequency and the highest proportion. Simply put, the word *Linguistics* occurred in 404 articles, which is the highest. This in turn implies that authors preferred using the word *Linguistics* to using the other words. It should be pointed out, on the other hand, that the word *linguistic* was the second most frequently used one. That is to say, it appeared in 354 articles, which in turn indicates that the word *linguistic* was the second most preferred one. Quite interestingly, the word *study* occurred in 319 articles (rank three). This in turn suggests that it was the third most preferred one for authors. With respect to the word *language*, it is interesting to point out that it ranks fourth. That is to say, it appeared in 283 articles. It can thus be inferred that the word *Linguistics* was the most preferred one for authors, followed by the word *linguistic*, the word *study*, and the word *language*, in that order. When it comes to the word *research*, it ranks sixth (the sixth highest). To be more specific, the word *research* appeared in 190 articles. With respect to the word *analysis*, it is interesting to note that it occurred in 158 articles and that it was the eighth most widely used one. Talking about the word *theory*, it appeared in 124 articles and was the tenth most frequently used one. Additionally, it must be pointed out that the word *structure* occurred in 77 articles and it was the twenty-fifth most widely used one. Finally, it should be mentioned that the word *feature* appeared in 39 articles and was the ninetieth most preferred one for authors. We thus conclude that the word *Linguistics* was the most preferred one for authors. Now the following graph briefly shows degree:

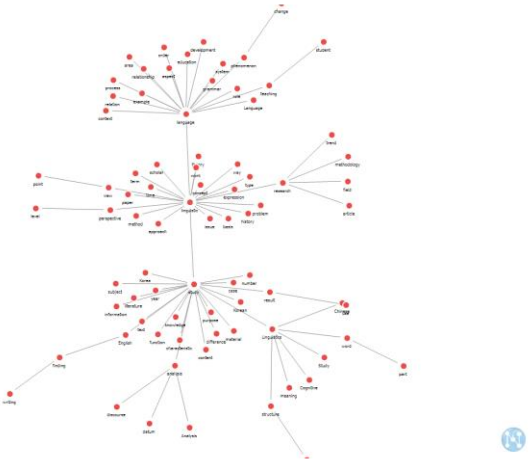
**Figure 2 Degree**



### 3.5. The Visualization of Words

In the following, we provide the visualization of words neighboring with the word *linguistics*. Figure 3 shows words neighboring with *linguistics* and links between them:

**Figure 3 Visualization of words neighboring with *linguistics***



As can be seen from Figure 3, the word *Linguistics* is indirectly linked to the words *linguistic*, *language* and *study*. Most importantly, the words *term*, *type*, *research*, *concept*, *method*, *level*, *issue*, etc. are directly linked to the word *linguistic*. This visualization shows us the picture of which neighboring words are linked to a keyword. For the visualization of synonyms, see Kang (2022a, 2022b, 2022c, 2022d). Quite interestingly, the words *process*, *aspect*, *teaching*, *education*, *example*, *context*, etc. are directly linked to the word *language*. These words are main words neighboring with the keyword *language*. More importantly, the words *text*, *subject*, *knowledge*, *function*, *result*, *difference*, *content*, etc. are directly linked to the word *study*. Such words are words neighboring with the keyword *study*. Finally, the words *structure*, *meaning*, *word*, etc. are directly linked to the word *Linguistics*. To sum up, this visualization provides us with the links between keywords and their neighboring words.

## 4. Conclusion

To sum up, we have analyzed 576 KCI articles published from 2002 to 2022. In section 3.2, we have argued that the word *linguistic* was the most widely used by authors, followed by the word *language*, the word *study*, the word *Linguistics*, the word *research*, and the word *paper*, in that order. In section 3.3, we have maintained that topic 7 was the most preferred one for authors, followed by topic 8 (topic 10), topic 5, and topic 9, in descending order. We have also maintained that as the 1st keyword, the word *linguistic* was the most preferred one. In section 3.4, we have contended that the word *Linguistics* was the most preferred one for authors, followed by the word *linguistic*, the word *study*, and the word *language*. In the case of the word *Linguistics*, it occurred in 404 articles, which is the highest. This in turn implies that authors preferred using the word *Linguistics* to using the other words. In section 3.5, we have provided the visualization of words neighboring with the word *linguistics*.

## References

1. Kang, N. (2022a). A Comparative Analysis of Search for and Look for in Four Corpora. *Advances in Social Sciences Research Journal* 9 (3): 168-178.
2. Kang, N. (2022b). A Comparative Analysis of Impressed by and Impressed within Two Corpora. *Theory and Practice in Language Studies* 12 (5): 819-827.
3. Kang, N. (2022c). On Speak to and Talk to A Corpora-based Analysis. *Theory and Practice in Language Studies* 12 (7):1262-1270.
4. Kang, N. (2022d). On Speak with and Talk with: A Corpora-based Analysis. *International Journal of Social Science and Human Research* 5 (8): 3354-3360.