Pakistan Social Sciences Review
www.pssr.org.pk

RESEARCH PAPER

The Productivity of Urdu Affixes in Newspapers: A Corpus Driven Research

Haroon Shafique [1] Dr. Muhammad Shahbaz [2] Ishtiaq Ahmad [3]

1. Lecturer University of Lahore, Gujrat Campus
2. Assistant Professor, Department of English, GC Women University, Sialkot
3. Ph. D Scholar Department of Applied Linguistics GC University, Faisalabad

PAPER INFO	ABSTRACT
	Affixation is one of the ways to form new words in Urdu through the use of prefixes, interfixes, infixes and suffixes. The current study introduces a new way to identify the potential affixes or morphological productivity of affixes in Urdu corpus which are used to form new words. The focus of the research is to find the productivity of affixes in Urdu. A huge corpus is compiled by the researcher to find the productivity of the affixes i.e. prefixes and suffixes. The Corpus driven approach is used by the researcher to analyze the affixes present in the corpus. Along with corpus driven approach, qualitative descriptive method is also employed by the researcher to examine the function of the quantified affixes from the corpus for the analysis of the data. The quantified results of the corpus show that بے is the most productive prefix while ی is the most productive suffix in Urdu. The results also reveal that suffixes are more productive than prefixes. The descriptive qualitative analysis shows that ی plays various functions in the process of affixation that is why it is the most productive affix in Urdu.

Introduction

Morphology is a discipline that talks about internal structure of words. It is a concept of a branch of linguistics in which different forms of words and their distinct use is investigated (Matthew, 1991).

Further, morpheme, lexeme and affixes are part of morphology and out of all, morpheme is the smallest unit of meaning, abstract vocabulary item is lexeme whereas affix is that morpheme which attaches to some other morpheme or in other words, it is a morpheme which is a stem, root or base. According to Katamba (1994), affixes can be attached before or after the base which are called either prefixes or suffixes. Moreover, morphological productivity is a core element in words formation, but it also challenges the uniform, solid description (Aronoff, 1976; Bauer, 2001; Plag, 1999). It is a common observation that new words formation in daily routine conversation is plentiful e.g. a person who is being gossiped about may be referred to as a gossipee, or a used book may be cleanish. This criterion of word formation is very productive in English, however, the degree of productivity for different affixes is different (Aronoff, 1976; Bauer, 2001) e.g. –ness is more productive (cleanness) as compared to –ity (cleanity).

The phenomenon of morphological productivity is complicated, and according to Bauer (2001), there is no consensus on a single notion regarding this productivity. For Robins (1990), morphological productivity is connected to something that has been recognized as a fundamental property of language and it is speakers' ability to create infinitely many new combinations out of the finite linguistic resources they have at their disposal. According to Plag (2006), in word formation process, it is speaker's ability to produce new words based on already existing words and with the aid of different recommendations. If a particular suffix (e.g. –ness) is used by speakers for making new words, then this process would be called productive.

New words production by using a particular affix in a notable frequency is called morphological production. In other words, it is a probabilistic continuum expecting the practice of potential words. At one pole of this continuum, unproductive affixes are present e.g. (th-truth) in English whereas on the other pole, inflectional suffixes are very much productive e.g. (ed, ing, s). These inflectional suffixes are used whenever there is grammatical conditions are suitable and no other blocking irregular form exists. Moreover, some other derivational suffixes (-ness and –attion) are very productive. In between these two extremes of poles, there are some less productive

derivational suffixes are present e.g. –ty. For some linguists, this concept is absolute – a configuration which is either productive or unproductive but these affixes keep on changing their placement on the scale of productivity continuum. Productivity of affixes cannot only be dealt via quantitative approach rather some qualitative morphological factors are also very much concerned e.g. rival affixes (convert adjective into nouns) - similar in their syntactic and semantic conditions.

Varying degrees of affixation productivity and its different measures have been projected in literature (Aronoff, 1976; Baayen, 1992, 2001; Bauer, 2001; Plag, 1999) whereas assessing this productivity is not an easy task (Bauer, 2001), but still there is a consensus that picking up the new words is mandatory in assessing the productivity although researchers are in trouble to explain what the actual 'new' word is? For determining this productivity, the most valuable procedure in previous studies is 'corpus-based approach' (Baayen, 1992, 2001).

Frequency is contended to be a significant parameter for the evaluation of morphological productivity and indispensable of new words formation but still there is a scarcity of research in this regard. The reason for this gap is due to the underestimation of quantitative approach in theoretical linguistics and it is somewhat relevant to the redundancy of empirical investigation because it is performance oriented instead of competence (Dressler & Ladányi 2000). Recently, a little attention has drawn on the expediency of corpora for quantitative investigation. The increase of linguistic material on electronic support and its easy access surely represent an important source to be exploited more and more in the future (see also Rainer 2003). By relying on literature, this study particularly focuses on corpus-based approach and presents a new method of assessing 'new' words and their productivity in Urdu affixes.

Urdu is an Indo- Aryan language and has the status of national language of Pakistan; moreover, it is one of the 23 national languages of India. In Pakistan, English is learned as second language after local language. It leads to variation in phonology and grammar causing problems for linguists. In local language, gender is assigned differently e.g. yogurt is considered as masculine in Urdu while it is feminine in other local languages. An additional complication arises

from the emigration, half a century ago, of the *Muhajirs*, the Muslims who left what is now India to settle in Pakistan after Partition. Due to this migration, different dialects are produced and they influence on other local languages. Affixation in Urdu is very frequent and especially its nouns are inflected gender, number and person via affixation e.g. gender in Urdu is categorized by four categories; marked and unmarked masculine (e.g. / dādā /and /g ar/), and marked and unmarked feminine (e.g. /larkī/ and /kitāb/). It is common observation that Urdu masculine nouns ends on these sounds /ā/, /ayā/, or, rarely, /ām / whereas feminine marked nouns use these endings / ī/ or /„ m /iyā/. These four categories are inflectional classes which are particularized by their endings (affixation) but not always e.g. /hayā/ shyness is unmarked noun for feminine but it ends on / ā/ which is marked ending for masculine noun. This research is focusing on these affixes and their productivity for 'new' word formation by using corpus-driven approach. Further, in the contemporary theoretical debate, the concept of frequency has gained a central position for encountering the processes of words formation. That is the reason, now it is requisite for a methodical investigation to focus on frequency of single affixes used in Urdu corpus. As the availability of electronic corpora has been increased, so, this research is immensely assisted.

The statement of the Problem

It is a fact that concordance is regarded as the parameters to evaluate the productivity but unfortunately very few studies deal with the productivity of affixes (Thornton, 1997). There is hardly any systematic study that deals with the aspect of productivity of affixes in Urdu morphology despite the fact that frequency assesses the word formation processes. A few studies are available that deal with productivity in different languages but this gap in Urdu is probably because of the underestimation of the useful tool i.e. corpus in the field of morphology. The advancement of electronics and technology has provided a support to the linguistics material to be analyzed quantitatively to explain the patterns of language. This study is an attempt to explore the productivity of Urdu affixes.

Limitations

This corpus-driven study has some limitations. The aim of this research is to find the most productive affixes. The use of affixes can also be situation based as the tenure in which the data is taken carries the prefix   prominent in the words like   . The word is predominant in the news of that time as the membership of the prime minister was challenged and was disqualified. Moreover, the software is certain cases does not distinguish between an affix and an orthographic form. So, the extracted data is analyzed manually by the researcher.

Delimitations

The corpus size is delimited by the researcher due to the time constraints. The data for corpus building is taken from January 2016 to January 2018. The researcher has delimited the corpus selection only to columns and editorials. As far as the affixes are concerned, the research only analyzes prefixes and suffixes because interfix and infixes cannot by analyzed by using corpus tools. Additionally, only five prefixes and suffixes are analyzed by the researcher. Moreover, the data is taken only from four newspaper by employing the criteria suggested by Nwogu (1997) and circulation. Additionally, the researcher has taken the affixes with concordance more than 10 because the concordance 10 or less than ten is not significant for 10 million corpuses.

Literature Review

According to Baayen (1992), there are numerous measures i.e. corpus bases, corpus driven to gauge the productivity and its aspects. Several studies have addressed the research aspect of exploring productivity in English as well in other languages. The study of Azmi (2013) is important in this regard who explored the productivity of English affixes in Magazines. The objective of the research was to explore the kinds of affixes and their grammatical forms in magazines. The data was analyzed by employing descriptive qualitative method which is suitable for the analysis of the corpus. This method helps to offer description analysis and classification of the different aspects of the phenomenon. The study revealed that there are five productive prefixes i.e. 're', 'un', 'dis', 'pre' and 'in' that are attached to the nouns. Moreover, twenty-five productive suffixes are revealed which either attach to a base or attach to a verb to make nouns. For instance,

develop is a verb and the addition of the productive suffix 'er' makes the verb a noun. The research significant as it classifies the affixes according to their forms but the data taken for the analysis is not huge enough to be called as representative data. This can be the criticism on the corpus size of the research.

Another study conducted by Nazean (2015) explored the productivity of 'ise' suffix in the published articles of medical articles. It was hypothesized by the researcher that 'ise' suffix is used productively in the field of medical sciences where a noun is converted into a verb with the help of the suffix 'ise'. The researcher analyzed 26 research articles from the field of medical science and concluded that 'ise' is a verbal suffix. The data revealed how vacuole and vascular i.e. nouns are converted into verbs vacuolise and vascularise with the attachment of 'ise' respectively. The research is undoubtedly significant as it proves the hypothesis of the researcher. However, the corpus of 26 articles is too little to analyse a linguistic feature.

The study by Saily (2008) analyzed the readiness of the suffixes with which they are attached to make new words or combinations (Bolinger, 1948). It was hypothesized that 'ity' is less productive as it is a foreign suffix which entered into English because of the result of language contact and ness is more productive because it is a native suffix. The corpus results proved the hypothesis to be true as the corpus showed 'ness' to be more productive as compared to 'ity'. The suffix 'ness' was used to convert adjectives into abstract nouns. The results also inferred that women use the suffix 'ness' more than men 'ity' which suggests that women have less interest in formal writing.

Liu and Shen (2012) carried out a corpus-based research to analyze the productivity and function of '-esque' suffix. The statistical results revealed that nouns, proper nouns and adjectives accommodate the affix '-esque to make new words. The new words with the suffix -esque are a feature of formal writing. Thus, the very affix is found to be more productive in formal writing. Additionally, -esque is found more in magazines, newspapers and academic writings. The researchers took the data from single corpus COCA but the same research an also be conducted on British National Corpus (BNS) to see the productivity of the affix. The productivity of the

suffix -esque can also be compared regarding two different corpora or two or more genres.

Other than English, a very few studies dealt with the phenomenon of morphological productivity in other languages. One of the examples is the analysis of the affixes in Italian. Ricca (2003) in a study analyzed 75 million newspaper corpus to see the productivity of the affixes. The findings of the research concluded that 'ri' and 'in' are the most productive affixes in Italian. If the results of morphological productivity are compared to the productivity of affixes in English, the derivatives are more productive in English and many other languages. However, in Italian, prefixes 'ri' and 'in' are more productive. So, prefixes play vital role for making new words in Italian.

The literature on morphological productivity not only addresses the productivity but it also suggests how the learners of English and second language learners take benefit from the productive affixes for learning. Buddingh (2005) carried out a research in a school and analyzed the language skills of the students of controlled and experimental group. The students of experimental group were given instructions for three weeks for twenty minutes per session and the results were compared. The results inferred that the students of experimental group were able to use more affixes in their speaking and writing. This study can also be helpful for the second language learners where the prescriptive rules of affixation help the learners to learn the second language more effectively.

Research Methodology

The study is conducted by using descriptive qualitative methodology to analyze the data. The data is collected with the help of corpus driven approach and then it is analyzed and classified into different categories and conclusions are drawn with help of descriptive qualitative analysis. Hence, two approaches i.e. corpus driven approach and descriptive qualitative approach help to collect the data and classify with help of analysis respectively. According to Maleong (1989) descriptive qualitative approach is an approach by which the researcher analyzes the data descriptively and classifies it.

Sampling

This corpus-driven research comprises of newspaper discourse from four Urdu Newspaper i.e. Nawa-i-Waqt, daily express, roznama duniya and Jang. For the development of corpus, only columns and editorials have been taken from these three newspapers. The data is taken from January 2016 to January 2018.

Rationale for Sample Selection

The criteria for sample selection are representation, reputation and accessibility and circulation that are proposed by Nwogu (1997). The forth criterion is added by the researcher i.e. circulation. According to Nwogu (1997), the selected corpus for research should be the representative corpus of the whole. In this regard, the journalistic data which is selected for the research is representative of Pakistani society as the newspapers are published from different cities of Pakistan namely Karachi, Islamabad, Lahore, Quetta, Sialkot, Gujranwala, Faisalabad, Rawalpindi, Peshawar and Sargodha. Therefore, the sample selected for the corpus is representative. For the criterion of reputation, the data for corpus building is selected from two renowned Pakistani English newspapers with huge number of readers and viewership. For the criterion of reputation, the journalistic data is written by well-educated people and edited by learned editors. So, the criterion of reputation is fulfilled. The criterion of accessibility suggests that whether the selected sample can be mechanized for research analysis or not. In this respect, the chosen sample is taken by coping from the official websites of selected newspapers which can be analyzed by software. As far as criterion of circulation is concerned, the selected newspapers are the most circulated newspapers of Pakistan.

Corpus Building

For the purpose of finding productivity of Urdu affixes, the data for corpus building are taken from the four Urdu newspapers. The corpus is built by using four sets and each set includes the data from one newspaper. The corpus building data is taken from 1st January 2016 to January 2018 with two years tenure. The columns and editorials of the newspapers are used for corpus building. The corpus

which is taken from Urdu newspaper consists more than 10 million words

Data Screening

Data screening is one of the important steps of corpus building. As in the world of information technology there are different ads and commercials which are the part of newspaper websites. In this research, the corpus is built by copying the data from official websites of selected newspapers. During the corpus building, different ads are omitted for the readability and analysis of the corpus by Antconc and lanxbox softwares.

Text Formatting

The copied data is transferred into "html" form from the websites of selected newspapers. Then copied data is pasted in the notepad files. For making the data readable by the software, it is converted into text files by using the Notepad standard settings.

Research Tool

Antconc 3.2.4w (2015) and lancxBox are the research tools for the corpus analysis. AntConc is a toolkit which is designed by Laurence Anthony for the analysis of corpus (Anthony, 2005). This is an extensive research tool which includes a concordance, words frequency generator, keywords frequency generator, function for cluster analysis and bundle analysis (Antony, 2005).

To find out the concordances and functionality of required words, Antconc is used. This research tool only analyzes text notepad files of the required text. Before adding the data in the software, the corpus files are converted into text files. Another tool, Urdu dost keyboard is used by the researcher to type the affixes in search bar of AntConc and LancsBox.

Corpus Driven Approach

To find the productivity of Urdu affixes, corpus driven approach is applied to find the productivity. Unlike corpus-based research, the research does not initiate the extraction of data from corpus having a theory, assumption or framework. The corpus is analyzed by the researcher and the productive affixes are quantified

with help of world list, advance features of world list and by typing different affixes present in the corpus.

Data Analysis and Discussion

For the data analysis, the researcher extracted prefixes and suffixes available in the corpus by using corpus driven approach and then the data is analyzed with help of descriptive qualitative methodology which deals with the classification of the data and then analysis. The prefixes and suffixes are quantified in the first stage of the analysis. The productivity of the affixes is analyzed through corpus driven approach. The classification of the productivity and with respect to the functions of the affixes is done through descriptive qualitative analysis.

Table 1: The frequency of affixes in Urdu

Prefixes	Concordance	Suffixes	Concordance
	840	ی	2451
	623		1972
	453		1283
	367	یں	751
کم	298		223
Total prefixes	**2581**	**Total suffixes**	**6680**

The functions of the suffixes
ی Suffix.

The corpus of ten million words through corpus driven approach reveals that ی is the most productive suffix in Urdu. The concordance of ی suffix in the data shows that it is the most productive suffix in Urdu. The corpus results reveal that ی is not only the most productive suffix but also among all the affixes i.e. prefixes and suffixes, it is the most productive affix. The concordance of ی is 2541. However, it plays many morphological functions when it is attached to the base. For instance, it serves as a gender, abstract noun, adjectival adverbial and diminutive marker.

ی as a Gender Marker

The given examples بچی and ننھی have been taken from the newspaper corpus where ی acts as a gender marker. Here   and ننھا

are root words while ی works as a gender marker when it is attached to the root words. It can be seen that ی changes the gender of the root word. The below given example is extracted from corpus which shows how ی works as a gender marker. The suffix ی does not always serves as a gender marker which can be seen in for the instance of کرسی Here, it is already a feminine noun.

ننھی بچی اپنے ارد گرد سے بے خبر تھی کہ ---

ی as an Abstract Noun Marker

ی which is the most productive affix also functions as abstract noun marker when it is attached to the base word. In the first example حکمران is a common noun and the suffix ی changes it into an abstract noun after the process of affixation. The other instances taken from the corpus are بادشاہی and نا انصافی .

حکمران اپنی **حکمرانی** کے نشے میں عوام کو کسی ----

ی as an Adjectival Marker

In the given examples, ادبی and اصولی the suffix ی is added to the roots which are nouns and are converted into adjectives after the morphological affixation respectively. The derived words are adjectives. Here ی serves as an adjectival suffix as it converts nouns into adjectives. The below given is the example where the suffix ی functions as an adjectival marker.

یہ تو سم کو معلوم ہو گا کہ یہ **ادبی** محفل کو زیب----

ی as an adverbial marker

ی suffix also functions as an adverbial marker. The corpus results show a few instances where ی makes the root adverb after suffixation. Here, فوری and جلدی are the two adverbs where ی attaches to the root and functions as an adverbial marker.

انتضامیہ کو فوری حکم نامہ جاری کرنے کی ضرورت ----

ی as Diminutive Marker

The given instance is a diminutive which is the result of affixation in the root. The root word is   While ی is a suffix. The addition of ی converts the noun   into جلسی which is a diminutive.

Unfortunately, there are very rare examples in the corpus where ی works as a diminutive marker. However, we can see many nouns which become diminutives because of the affixation of ی. For Example, پنکھی ، میزی

جبکہ مخالفین کے خیال میں یہ ایک **جلسی** سے زیادہ۔۔۔۔۔

as plural marker

The corpus shows that   is the second most productive affix in Urdu. The suffix   is found to be very productive with 1972 concordance in 10 million words. The suffix   works as a plural marker in the given examples. The root word کمرہ is changed into کمرے after the addition of   as a suffix. The suffix   also works as oblique marker. However, the corpus needs to be analyzed manually to see the function of the suffixes. For instance, the suffix   in کمرے serves as a plural as well as oblique marker in two different cases in corpus. It is also evident in the case of the derived word   where   is again a plural as well as oblique marker.

آفس کے تمام **کمرے** نہایت قیمتی فرنیچر سے لدے ------

as plural marker

The suffix   proves to be third most productive suffix in Urdu. The concordance of the very suffix is 1283. It is noteworthy that the most productive suffix has more than double concordance than that of   . The suffix   in the given examples پہاڑوں and   added in the nouns پہاڑ and   The grammatical category of the root remains the same after affixation. Here,   is a plural marker which is third most productive marker in Urdu.

پاکستان بلا شبہ **پہاڑوں**   **وادیوں** کی جنت ہے۔۔۔

یں as Plural marker

After the suffix   the fourth most productive suffix is یں which is a plural marker with 751 frequency in 10 million words. The given examples حرکتیں and سرنگیں are the derived words from the roots حرکتand   respectively. The example given below is taken from corpus which shows how یں works as a plural marker.

سفر کے باد معلوم ہوا کہ یہ **سرنگیں** چائنہ اور پاک ارمی نے -----

as a plural marker

The corpus results exhibit that the suffix   is also a productive suffix which is used to make new words by pluralization.   is a Persian affix that has been added into Urdu for making plural. The concordance of the suffix at is 223. The examples of the derived words are taken from the corpus. For instance,   is a root whereas   is a derived word because of the affixation of the suffix   which makes the root word plural.

اگر ملک کے   یہی رہے تو بہت جلد مزید قرضے ۔۔۔۔

The Function of Prefixes

as a Negative Marker

The given examples   and   are the derived words after of process of affixation. Here   works as a prefix which is also the most productive prefix in Urdu. The frequency of   as a prefix is 840 in ten million words.   is a noun, also works as a verb and depends upon its contextual use while   is a negative marker which works as a prefix. The derived word is بے قدری which is a noun meaning by being not careful.

اس نو جوان کی یہ حرکت ھی اس کی **بے قدری** کا منہ بولتا۔۔۔۔۔

as a Negative Marker

The corpus results show that   is a prefix with second highest concordance in corpus. The concordance of the prefix   is 623. The instances نا اہل and نا کامی are the derived words in which اہل and کام are the root words. Here,   is a negative marker which has a native Urdu origin. The addition of   in the process of affixation adds negative sense. For Example, کام means work while ناکامی means the state when someone is unable to do work or someone is failed to do work. The root کام is a noun while ناکامی is also a noun.

سپریم کورٹ کے فیصلے کے نتیجے میں ن لیگ کے قاید کو **نا اہل** ۔۔۔۔

as a Negative Marker

The prefix　 is an Arabic negative marker which is also used in Urdu. The concordance of the prefix shows it is widely used in Urdu language as well. However, it is not used independently rather it is used as a prefix. Here, the example shows that　　 is an adjective.　 is a negative marker here that is attached to the root word i.e.　　.

انسانی جسم مین موجود یہ　　　　سیلزبھت سے کام ------

as a possessive marker

　　　is a Persian marker that is used to show the possession of some state or quality. The corpus reveals that　 is used significantly in Urdu. The given example　　　　means the one who enjoys power or government. Here, the prefix　 works as a possessive marker which adds the meaning of the root. The other extracted words from the corpus having as a prefix　 are　　　　　.

در اصل حقیقت یہ ھے کہ پاکستان کا ہر　　　　شخص قانون -----

کم as a negative marker

The prefix کم is also a negative marker. The frequency of the prefix km is 298. In the given example کمبخت،　　 is a root word which is a noun. The addition of prefix کم changes its meaning. The prefix کم is thus a negative marker. The other examples extracted from the corpus are کم حوصلہ کم ذوق.

مہنگائ کی ماری **کمبخت** عوام جائے تو کھاں جائے۔۔ ---

Conclusion

The study focuses on the morphological productivity which helps an affix to be used for the production of new words. The corpus has been compiled by the researcher to examine the morphological productivity of affixes in Urdu. The data i.e. morphologically productive affixes are extracted by the researcher for the quantification by employing corpus driven approach. Further, the data is analyzed descriptively and classified into different categories using descriptive qualitative method. The findings of the research reveal many facts. The prefix　　 is the most productive prefix with highest concordance in the corpus while the suffix ی is the most productive suffix in Urdu. The descriptive qualitative analysis shows

that the affixes play different roles in the process of affixation. The statistical analysis of the corpus shows that suffixes are more productive as compared to the prefixes which is evident through the statistical difference in the use of prefixes and suffixes, however this productivity of suffixes is because of their multifunctionality. The suffix ی functions as a gender marker, abstract noun marker, diminutive marker, adjectival marker and adverbial marker. The morphological productivity of ی is because of its ease to attach root words to derive new words. For the future researchers, it is suggested that the same study can be conducted by extending the word tokens or the corpus size. Additionally, the corpus for future researches can also be taken from different genres.

References

Anthony, L. (2005, July). AntConc: design and development of a freeware corpus analysis toolkit for the technical writing classroom. In *Professional Communication Conference, 2005. IPCC 2005. Proceedings. International* (pp. 729-737). IEEE.

Aronoff, C. E. (1976). Predictors of success in placing releases in newspapers. *Public Relations Review*, 2(4), 43-57.

Aronoff, M., & Anshen, F. (2017). Morphology and the lexicon: Lexicalization and productivity. *The handbook of morphology*, 237-247.

Baldry, A., & Thibault, P. J. (2005). *Multimodal transcription and text analysis*. London: Equinox.

Bauer, L., & Renouf, A. (2001). A corpus-based study of compounding in English. *Journal of English Linguistics*, 29(2), 101-123.

Baayen, H. (1992). Statistical models for word frequency distributions: A linguistic evaluation. *Computers and the Humanities*, 26(5-6), 347-363.

Baayen, H., & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. *Linguistics*, 29(5), 801-844.

Bolinger, D. L. (1948). On defining the morpheme. *Word*, 4(1), 18-23.Shen, M., Liu, D. R., & Huang, Y. S. (2012). Extracting semantic relations to enrich domain ontologies. *Journal of Intelligent Information Systems*, 39(3), 749-761.

Buddingh, M. (2010). *The effects of teaching roots and affixes on the vocabulary development of underperforming students* (Doctoral dissertation).

Dressler, W. U., & Ladányi, M. (2000). On contrastive word-formation: German and Hungarian denominal adjective formation. *Trends In Linguistics Studies and MonographS*, 130, 59-74.

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be?. *Applied linguistics*, 11(4), 341-363.

Gale, W. A., & Church, K. W. (1991). Identifying word correspondences in parallel texts. In *Speech and Natural Language: Proceedings of a Workshop Held at Pacific Grove, California, February 19-22, 1991*.

Krott, A., Baayen, R. H., & Schreuder, R. (2001). Analogy in morphology: modeling the choice of linking morphemes in Dutch. *Linguistics*, *39*(1; ISSU 371), 51-94.

Katamba, F.1993. *Morphology"Morpheme the Smallest Unit of Meaning"*. The Macmillan press Ltd. Hound mills: United Kingdom.

Liu, W., & Shen, H. (2012). A Corpus-based Analysis of English Suffix--esque. *Theory & Practice in Language Studies*, 2(4).

Maleong, L. 1989. *Metodologi Penelitian Kualitatif*. Jakarta: Depdikbud.

McMahon, A. M., & April, M. (1994). *Understanding language change*. Cambridge university press.

Myerson, R. F. (1976). *A study of children's knowledge of certain word formation rules and the relationship of this knowledge to various forms of reading achievement* (Doctoral dissertation, Harvard University).

Năznean, A. (2015). The productivity of the ise suffix in a corpus of medical articles. Studia Upetru *Maior- Philologia*, 19.

Nwogu, K. N. (1997). The medical research paper: Structure and functions. *English for specific purposes*, *16*(2), 119-138.

Plag, I., Dalton-Puffer, C., & Baayen, H. (1999). Morphological productivity across speech and writing. *English Language & Linguistics*, *3*(2), 209-228.

Plag, I. (2006). Productivity. *The handbook of English linguistics*, 537-556.

Ravosa, M. J. (1991). Structural allometry of the prosimian mandibular corpus and symphysis. *Journal of Human Evolution*, *20*(1), 3-20.

Säily, T. (2014). Sociolinguistic variation in English derivational productivity: Studies and methods in diachronic corpus linguistics. *Mémoires de la Société Néophilologique de Helsinki*.

Thornton, T. F. (1997). Anthropological studies of Native American place naming. *American Indian Quarterly*, *21*(2), 209-228.

Van Nice, K. Y., & Dietrich, R. (2003). Task sensitivity of animacy effects: Evidence from German picture descriptions. *Linguistics*, *41*(5; ISSU 387), 825-850.