Representation of Language in Sri Lankan English Newspapers: Key Word Analysis with reference to two English Newspapers

H.A.S.L. Amarasinghe
Institute of Technology, University of Moratuwa.
lakshikaa@itum.mrt.ac.lk/lakshikasandarani@gmail.com

Abstract

It is more often concerned by linguists that due to their physical absence and less voice before the public negatively impact to create a less recognition on Language and Linguistics related topics in newspapers even though, they are practiced in every aspect in the society. Considering a key word analysis employed in Corpus Linguistics, this study investigated to which extent Language and Linguistics related areas are represented in the News section of the Sri Lankan English newspapers. The sample included the selected news articles from the electronic versions of two newspapers that had been published for six months in 2021. The total searched word count of this corpus was 15636 which was analyzed for key words through Browserling online word frequency counter to investigate the occurrence of the terms <language>, <languages>, <linguistics> and <education> in the word list. This corpus-based methodology revealed that these key themes have not become interesting areas in Sri Lankan English newspapers. Similarly, the semantic fields that are related to these key themes are intertwined with current socio-political discussions in the country. Furthermore, the proper nouns like names of leading figures in different fields of the country deserve a prominent place in this media discourse as well. Overall results of this study indicate that the Language and Linguistics related themes are to be more positively addressed in news articles as they are more commonly practiced in authentic situations among the public. Also, more research has to be conducted related to corpus-linguistics by using the International Corpus of English – Sri Lanka (ICE-SL).

Key Words: *newspapers, Sri Lanka, key word analysis, language, linguistics*

Introduction

It is more often concerned by linguists about their physical absence and less voice before the public, influence to create certain misinterpretations of language in media, whenever aspects of language

science are discussed with general audience. "The Journal of Sociolinguistics" describes the views of researchers regarding misinterpretations of ideologies of the language and linguists via media. (Aitchinson, 2001; Milroy, 2001; Herler,1998). These research studies uplift the malfunctional relationship between language and media.

Media language has always attracted the attention of linguists, particularly Applied Linguists and Sociolinguists. It further reveals some practical reasons for this popularity of media in the field of Sociolinguistics; easy to access language data for research, access to a large component of written or spoken language data of the general public, it is interesting explore how media touches the demographic variations of the public and they act as the voice of different cultural, political and social views of the public. These points provide the evidence with the huge impact of media on the representation of language use in a selected context. The impact of it would create a large influence of the ideology of the general public, which truly negatively makes the linguists unhappy.

Newspapers play a major role in the Sri Lankan context keeping it as a powerful source of information eventhough there is a rapid development in terms of media. The readers still explore many in different fields in terms of locally and internationally, especially through English newspapers. In the recent past the globe experienced new areas in terms of education, health and so on with the outbreak of Covid-19 pandemic which deviated some significant topics and themes usually addressed through the newspapers. Linguistics and Language are two mighty concepts in the world which enjoys a power of creating a generational transience in the discipline of Linguistics. These concepts address the human issues in a genuine manner but they are not well highlighted in today's newspapers due to certain issues.

This crucial situation has been addressed in the sub disciplines of Linguistics called "Media Linguistics" as well as in Sociolinguistics"". Media linguistics studies how language is used in media and also it further reveals that, this discipline has its roots in conversational analysis in Sociolinguistics. Mass media texts such as excerpts from newspapers, television and radio become the target objects of this sub discipline. Moreover, the Corpus Linguistics which is related to producing corpora in different sub fields relevant to studies. In these studies, mini-corpora are compiled and data are analyzed thorough software toolkits like Browsling and Antconc and WordSmith. This analysis can be done in ways like keyword analysis, concord and wordlists.

So that, the present study aims at discovering the representation of language in Sri Lankan English newspapers in the light of Media linguistics, Sociolinguistics and Corpus linguistics.

Research problem

Within Sri Lankan context, the subject area "Linguistics" is not that much taken into discussions on the stage of public relations at a satisfaction level. Many of the news articles in the news section are dedicated for the prominent political events while the Features display Art, Literature, History and other events. All these news productions are done in the light of language despite the fact that Linguistics and Language become least touched news areas in Sri Lankan English newspapers. Simultaneously, the writers do not show a positive tendency to express their views and preset study findings related to these two fields via Newspapers.

Within this context, it is necessary to intervene the way how the language related concepts or issues are represented through key words in Sri Lankan English newspapers. There are few published research articles on this area in Sri Lanka, keeping this research area still uninvestigated. So that this study investigates how Language and Linguistics related topics are represented in Si Lankan English newspapers.

Research Questions

1.   What are the key words occurring in the News sections of the English newspapers in Sri Lanka?
2.   To what extent Language related themes are discussed in   these newspapers?

Objectives of the study

-   To investigate the key words of language related news articles in Sri Lankan English newspapers.
-   To explore how language related themes are represented in the Sri Lankan English newspapers.

Literature Review

Key words are significant words which are related to the theme of a particular text in any code. They carry the meaning of it while embedding related concepts with the art of language. Moreover, the keywords in a text display the weightage given for each sub theme of the particular text. So

that, investigating the keywords in a particular text indicated the popularity of the particular area among the public approximately. This research area can be explored in the light of Corpus Linguistics which refers to compiling a database of a particular area exploit it in different ways to in explore answers to research problems and compare it with a larger existing corpus. Choe (2017) has conducted a research study on "Relating Key words to the "Top ten news of the Year" in Korean Newspapers, explores the relationship between the mechanically extracted words and the top ten news of the year". This research study employs the Dunning's Log Likelihood (LL) and chi-square (Scott & Tribble, 2006) which reveals 100key words in different point of views.

Corpus Linguistics

Corpus Linguistics quantitatively analyses corpora up to several million words of spoken and written language enabling the researchers to establish frequencies and probabilities of words and phrases often tighter with demographic characteristics of their users (Paltridge & Phakiti, 2015). This suggests that Corpus linguistics has become a key mode of linguistic inquiry throughout the history. Baker (2004) explains how a large compilation of electronic texts which is carefully sampled, contributes to the linguistic researchers to study the language representation of a particular variety. This broad area has been studied by a number of researchers: Corpus Studies of Lexical Semantics (Stubbs, 2001), Corpora and Applied Linguistics (Hunston, 2002), and Using Corpora in Discourse Analysis (Baker, 2006). A number of diachronic studies have explored many possible specific areas of Corpus linguistics.

Corpus Linguistic Studies in South Asia

The varieties in English in different parts of the world and their unique features which are bound with the demographical varieties have been addressed by a number of linguists and researchers. In his book: "Asian English Beyond the Canon", Kachru (2005) deploys the term "Asian Englishes" to introduce the varieties of English existing in South Asia. His robust claim, "The English language is discussed as a language *in* Asia, but not *of* Asia" (Kachru, 2005), opens up a new arena to the status of English in South Asia. It gives a more weight to the compilation of ICE versions of South Asian countries. It is a recorded electronic sample of written and spoken language varieties of a language used in a particular country. These

electronic documents also reflect a blend of Asian culture in different varieties of a foreign tongue (English) which provide the future researchers with evidence to explore specific areas in Corpus Linguistics as well as Sociolinguistics.

Corpus Linguistic studies related to English newspapers in Sri Lanka

This target area of Linguistics still has not been widely explored by the researchers in Sri Lanka. Only a limited number of published research papers are available in this field. Hettiachchi & Ranaweera (2012) conclude their research on "A Corpus Survey of Sri Lankan English in Selected Sections of Sri Lankan English Newspapers of Sri Lanka and the Implications for Writing Authentic Lesson Material in Tertiary level ESL Context", with the finding of Sri Lanka lexical items are present in Sri Lankan English newspapers which is contradictory to previous Studies"

International Corpus of English in Sri Lanka (ICE-SL)

The Sri Lankan component of ICE is a paramount importance with regard to empirical research into new Englishes (Bernaisch et al., 2020). It is comprised with two parts: the written part of ICE-SL and the spoken part of ICE-SL, after a collective effort of a collaborative team work of academics of the University of Giessen and the University of Colombo.

Present Study

This study adopts a corpus-based approach to explore the language construction in Sri Lanka English newspapers by investigating the key words of newspaper articles in News section.

Methodology
Approach

This research study focuses on investigating how language related issues are represented in Sri Lankan English newspapers through a key word analysis. As this intervention explores the frequency of the occurrence of language related words as well as semantic fields related to the key words this study can be identified as a mixed method approach. It focuses on investigating both qualitative and quantitative data.

Sample

The present study employs the electronic versions of two of the English newspapers in Sri Lanka "Sunday Observer". There are some

background reasons to select these particular English newspapers because they can be easily accessed through the archives under each section is the first and foremost reason behind this. Moreover, another reason is that the electronic versions of these newspapers are available in online and they are popular among the readers as well and have been used for research purposes even among the lowest number of published research studies related to English newspapers in Sri Lanka.

The News section was selected as the target area of this study because it comprises news related to the general public which is not limited only to a set of audience. The other news sections such as Features and Sports approximately target a particular audience as well as authors which limits the discussion of the general audience. As this research study focuses on the news of the general audience, it uses the News section as the target area. Moreover, this section is dedicated to the most up to date news items in the country which serves the value of the research as well.

News related to the any area related to language, linguistics and education was selected as the sample for the current research study. The selection was mainly based on the title of the news article as well as the content of the it They were extracted from the Sunday Observer and Sunday Times newspapers published from 01st 0f May to the end of October of 2021which covered a period of six months.

Data collection Instruments

The Sunday Observer and Sunday Times English newspapers from May to November 2021.

Procedure

This study is a corpus-based study regarding the language representation in Sri Lankan English newspapers. A corpus of texts collected from the Sunday Observer and Sunday Times e-newspapers which were published from May to October 2021 which covers a period of six months. Even though they are some more English e-newspapers freely available in the internet, the Sunday Observer and Sunday Times can be identified as popular English newspapers among the readers of Sri Lanka.

After selecting the sample for the research study, an electronic search was done to collect the relevant newspaper articles under three main themes "Language, Linguistics and Education". In some news articles these key words are there in the titles which the others imply the relevancy of the

articles to the themes in an indirect way. When the title of the research article and the contents displays any related above-mentioned topics were saved as a word documents. And then, they were converted in to electronic text files.

By collecting newspaper articles an electronic text file was created. After creating this sub corpus, the next stage of the research study was to identify the key words related to these terms. The key words are related to the frequency of occurrence of them in these particular texts.

Then, Browserling online word frequency counter was employed as the tool to create a list of key words in the generated electronic file. After that, a statistical analysis was done to investigate the occurrence of frequency of language related themes in selected texts.

Data Analysis

The frequencies of the occurrence of the node terms "language, linguistics and education "were measured through the Browserling online word frequency counter

Key words	Sunday Observer	Sunday Times	Total
<language>	-	9	9
<languages>	5	-	5
<linguistics>	-	-	-
<education>	33	73	106
Total number of search terms	6666	8970	15636

Table I: Occurrence of four terms across the two newspapers

The corpus generated by using the Sunday Observer newspaper had a word list of 6666 while the corpus of Sunday Times had 8970 words. Then each mini corpus was fed to Browserling online word frequency counter to generate the frequency list of the words. The analysis shows that the term/ node "language" had no any occurrence in the corpus created based on the articles in the News section of Sunday Observer. The articles were also selected based on the subject area of the news item. Additionally, the key word Languages has appeared in five times in this newspaper. This indicates that the language related area has been represented in to some extent in this newspaper. Next, this key word comes in 9 times in the same section of the Sunday Times newspaper during the selected period (from May – October

2021). This shows that the second newspaper also has given some palace in it to represent language related topics in the News section.

According to the data analysis, there has not been any discussion on Linguistics in the News section of in both newspaper indicating that this key theme has not become an interesting area in the print media. The writers of the articles who focus the general public through the news section seem not to consider Linguistics as an appealing topic to the readers. This further indicates that the area Linguistics is practically used in every aspect in the society in terms of Sociolinguistics, Applied Linguistics, Forensic Linguistics, Psycholinguistics and Media linguistics and so on the general public is not that much familiar with this subject area and its applications. This is kind of a negative remark for the linguists as well as students who study this area.

Next, the key word education has been occurred 33 times in the Sunday Observer and 73 times in Sunday Times showing that this field is addressed more frequently in the News section of the selected two newspapers. It marks the highest frequency in the sample of this research study. This became a very common topic in every media all over the world with the outbreak of Covi-19 pandemic situation because the education system underwent a drastic shift from many traditional methods to digital methods during this period.

Semantic fields	Types
Language	Book, Sinhala, Tamil, writers, publishing, reading, writers, Rotaract, Chinese, literature, edition, research, speakers, research, conference, faculty, college, departments, journal, lip-reading, Hindi
Education	University, student/s, schools, teacher/s, children, online, school/s, strike, salary, universities, principals, examination, youth, virtual, Kotalawala , educational , knowledge , studies , school children , ugc , teaching , lessons , higher, parents , study , professor , scholars , ministry, technology, graduate, curriculum, learning , allowance , sociology , primary , protest , Cambridge , examination , tuition , classroom , learning , syllabus , lectures, assessments , virtual
Health	Pandemic, Covid-19, people, vaccination, health, public, masks, spread, severe, donations, vaccine, AstraZeneca, doses, responsibility, booster,

Table II: Semantic fields and related types

The key words generated under this research study can be discussed in relation to the key semantic fields as in above. The lexical set in this chart belong to the same conceptual area. The first sematic field selected in this research study is Language. This semantic field is discussed articles in the News section of two newspapers in relation to the identified lexical terms. Next, the semantic field of Education is discussed with a large number of terms as in the chart because, many changes occurred in the field of education during the period of past one and half years. Especially the discussions online teaching. Kothalawala act and the protest of Teacher and principals influenced a lot in print media in the same period. This indicated how current issues are repeatedly reported in print medial while reducing other existing concepts in the society. Moreover, these social discussions are used for the popularity of politics also minimizing the use of newspapers for linguistics related research purposes.

Proper nouns

Both sub-corpora of compiled from News articles of two English newspapers reveal much more use of proper names related to the current themes in the society. Even though the news articles selected as the sample of the research through the topic of the news item according to the author's view, as today language, education and linguistics terms are intertwined with current issues in the country. Moreover, in any newspaper article corpora there are occurrences of the proper names of the leading political figures, government officers and social activists. However, the occurrence of proper names in the created two corpora are insignificant in frequency in comparison to the major key words.

Limitations

1. As the data for the research study was gathered during the pandemic situation, the expected data could not be obtained exactly. Many of the articles in the News section of both newspapers and generally in other media also were dedicated for Covid-19 related issues as well as much of the political issues as usual.
2. The time period taken for data collection was not enough due to the above-mentioned point. The research study would me more effective if the data collection was done for at least one year.
3. This corpus-based study is needed to be conducted by comparing the data with ICE-SL which could not be done due to time constraints.

However, the author intends to obtain the ICE-SL from the relevant scholars and extend this research in future.

Discussion

There is a need of a discussion on Language and Linguistics through newspaper in the Sri Lankan contexts as these areas are not positively addressed in print media. The linguists, researchers and the students who study these areas need to be pioneers in this task because the general public employs many more pragmatic functions of language and many sub-areas of Linguistics in authentic situations without any theoretical knowledge. This indirect highlight the descriptive Linguistics. The whole society is interconnected merely because of the functions of language and aspects of Linguistics. Therefore, the author intends to focus the attention of the above-mentioned parities to bring the topic before the public through print media. On the contrary, the data analysis indicates that some distant areas for the general public which are not practically used by the individuals are more frequently discussed through the newspapers due to political influence.

Future research

It is necessary to conduct more research on language representation in Sri Lankan English newspapers through key word analysis using a a large reference corpus. The gratitude should be paid for the scholars and professional who pioneered to compile both spoken and written corpora of Sri Lankan English. This data base is an assert for corpus-based studies in Sri Lanka. As an extension of the current research the author intends to obtain the ICE-SL from the relevant authorities to continue research on this area as this is still uninvestigated in the Sri Lankan research context.

References

Aitchison,J. (201). Misunderstanding about Language: A Historical overview. Journal of Sociolinguistics 5(4).611-619

Archive. Times Online - Daily Online Edition of The Sunday Times Sri Lanka. (n.d.). Retrieved December 19, 2021, from https://www.sundaytimes.lk/210103/archive/

Baker, P. (2004). Querying keywords. *Journal of English Linguistics*, *32*(4), 346–359. https://doi.org/10.1177/0075424204269894

Baker, P. (2006). Using corpora in discourse analysis. https://doi.org/10.5040/9781350933996

Bernardini, S., & Aston, G. (2002). Review of stubbs ((2001)): Words and phrases. corpus studies of lexical semantics. *International Journal of*

Corpus Linguistics, *7*(2), 283–295. https://doi.org/10.1075/ijcl.7.2.07ber

Browse by Week. Sunday Observer. (n.d.). Retrieved December 19, 2021, from https://www.sundayobserver.lk/date/2021-12-05

Ensslin,A. & Johnson, S. Langauge in the News: Some Reflections on Key Word Analysis Using WordSmith Tools and the BNC. academia.edu https://www. /276923/

Heller, M. (1999) Sociolinguistics and Public debate. Journal of Sociolinguistics 5(4). 626-631.

Hunston, S. (2002) Corpora in applied linguistics. Cambridge University Press

Jae-Woong Choe. (2017). Relating keywords to the 'Top Ten News of the year' in Korean newspapers. *Language Facts and Perspectives*, *41*(null), 113–138. https://doi.org/10.20988/lfp.2017.41..113

Kachru,B. (2005). Asian Englishes: beyond the canon. Hong Kong University Press.

Milroy,J. (2001) Response to Sally Johnson; Misunderstanding Language? Journal of Sociolinguistics 3 (2) 276-281.

Paltridge, B., & Phakiti, A. (2015). Research Methods in Applied Linguistics. Bloomsbury Academic.

Ranaweera, M., & Hettiarachchi, S. (1970, January 1). *A corpus survey of sri lankan English in elected sections of English newspapers of Sri Lanka and the implications for writing authentic lesson material in a tertiary level ESL context: Semantic scholar*. undefined. Retrieved December 19, 2021, from https://www.semanticscholar.org/paper/A-CORPUS-SURVEY-OF-SRI-LANKAN-ENGLISH-IN-SELECTED-A-Ranaweera-Hettiarachchi/f4da16572e050a3ab2bc1aaadb572ddc1e0703fb

Scott,M.& Tribble, C. 92006) textual Patterns: key Words and Corpus analysis in Language Education. Benjamins.