

Aspect in English and Mugali Rai: A Contrastive Study

Ichchha Purna Rai¹

timma.raai@gmail.com

¹Dhankuta Multiple Campus, Dhankuta, Tribhuvan University, Nepal

Abstract

The paper attempts to explore the aspect system in Mugali Rai, spoken in Dhankuta district of Nepal by a very few people and compares and contrasts Mugali Rai aspect system with that of English aspect system. There are only two aspects in Mugali Rai, namely, perfect and progressive. Perfect aspect can be categorised into past perfect and present perfect in terms of time dimensions. Similarly, progressive aspect is also categorised into past progressive and present progressive from time dimensions. All types of aspects in Mugali Rai are morphologically marked. On the contrary, aspect system in English is not only morphologically marked. There are several complex constructions using have+past participle, be+present participle, and have+been+present participle for perfect, progressive and perfect progressive, respectively. Mugali Rai has only four structures for aspect whereas English has 17 different types of aspectual structures described in examples (24-40). It is really a challenging job for Mugali Rai learners to conceptualise these different structures. Finally, this paper finds out EFL problems and suggests some pedagogical strategies for teaching and learning English aspect system as a foreign language to Mugali Rai learners.

Keywords: aspect, past progressive, present progressive, past perfect, present perfect

1. Introduction

Aspect is certainly different from tense though they are interrelated. There are mainly two types of aspect, namely, grammatical aspect and lexical aspect. In grammatical aspect, aspect is manifested through grammatical operations whereas lexical aspect refers to semantic properties of verbs whether or not an action is characterized by duration, an endpoint, or change (Cowan, 2009, p. 352). In this paper, grammatical aspect is described. Aspect expresses how a speaker views the action of the verb. An action, that is seen as bounded and complete, is perfect in aspect. If the action is seen as incomplete, it is imperfect in aspect; if seen as repeated, it is iterative; if seen as occurring regularly, it is habitual (Cowan, 2009, p. 251). In another definition, aspect describes the temporal shape of events or states (Payne, 2003, p. 238). Aspect is categorised into two types, namely, grammatical and lexical aspect. In grammatical aspect, aspect is manifested through grammatical operations, whereas lexical aspect refers to semantic properties of verbs whether or not an action is characterized by duration, an endpoint, or change (Cowan, 2009, p. 352). Similarly, aspect is defined through three dimensions, viz. perfectivity, which is categorised into two binary concepts such as perfective vs. imperfective, sequentiality, which is categorised into perfective vs. perfect, and immediacy, which is categorized into remote vs. vivid (Givón, 2001, p. 287).

1.1. Research Questions

There are a number of research issues about aspect in English and Mugali. This study mainly focuses on the following research questions.

- a) Is there perfective or perfect aspect in English and Mugali?
- b) How is aspect constructed in English and Mugali?
- c) What are the similarities and differences between English and Mugali aspect?

- d) What are the pedagogical implications?

1.2. Objectives of the Study

This study has the following objectives:

- a) to analyze the aspectual construction in English and Mugali Rai
- b) to contrast between English and Mugali Rai aspect
- c) to point out pedagogical implications

1.3. Significance of the Study

This study is limited to contrastive study between English and Mugali aspect. It mainly analyses aspect system in both languages. It focuses on analysing Mugali aspect to compare and contrast it with English. It would be significant for those who wish to study endangered languages of Nepal and compare and contrast these languages from pedagogical perspective. It would also be significant for those who have been teaching English as a foreign language across the world.

1.4. Methods of the study

This study is based on descriptive linguistic approach, specifically on Contrastive Analysis (James, 1980; Lado, 1957). Contrastive analysis between two languages is based on some previous works (Rai, 2007, 2012). In the case of sources of data, English data were taken from the secondary sources which include Givon (2001), Payne (2003), and Cowan (2009). On the contrary, Mugali Rai is undocumented and undescribed language. For the first time, I visited the field (Muga VDC of Dhankuta, East Nepal) and described it in my first fieldwork (Rai, 2011).

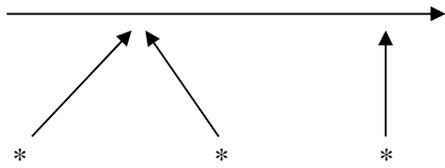
2. Analysis of English and Mugali Aspect

In this study, aspect in English and Mugali is described. There are different types of aspects expressed through the inflections in the verb such as perfect and perfective aspect which express the completeness of an action, imperfect aspect and progressive which express an ongoing activity or process, iterative aspect which expresses repetitiveness of an action, inceptive aspect which signals the beginning of an action, habitual aspect which expresses the action occurring regularly, inchoative aspect which signals entrance into a state, and lexical aspect.

2.1. English Aspect

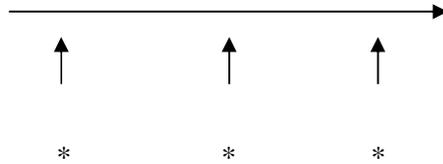
In English, they are represented by a number of ways. There are some puzzling concepts in different types of aspects. So, it is important to define some concepts of aspect found in the English language. In many traditional grammar books of English, perfect and perfective are treated as one and the same, but this is not the case. There are considerable differences between perfect aspect and perfective aspect, which can be presented through the following figure adapted from Givon (2001, p. 296). This figure can help distinguish perfective from perfect aspect more clearly.

a. Perfective Past



relevance time event time time of speech

b. Perfect Past



event relevance time of speech

Perfective refers to a situation which is seen as a whole, regardless of the time contrasts which may be a part of it. On the contrary, the perfect refers to a past situation where the event is seen as having some present relevance (Crystal, 2008, p. 356). Givon (2001, p. 296) has given the following examples to make the difference between past perfective and past perfect.

- (1) He came in and ate rice (past perfective).
- (2) He came in. He had (already) eaten rice (past perfect).

The difference between perfective and perfect may also be defined as the in-sequence and out-of-sequence. In (1), there is in-sequence between two clauses which expresses the perfective aspect, whereas in (2), there is out-of-sequence between two clauses which expresses the perfect aspect. If we look at these two examples mentioned in (1) and (2), there is no perfective aspect in English. In English, the aspectual meaning of perfective aspect is expressed by past tense. So, there is only perfect aspect but not perfective aspect in English.

In English, there are only two aspects, viz. perfect aspect and progressive aspect, which are represented in verbs. Two aspects are expressed through auxiliary verbs and the form of main verbs. The progressive aspect which represents an ongoing action is structured with *be+present participle (-ing)* and the perfect aspect which represents action that is complete is constructed with *have+past participle (-ed/-en)*. These two forms of aspect express the aspectual meanings in English. The progressive and perfect aspects are combined to form perfect progressive aspect in English. All of these aspects are combined with the three dimensions of times, viz. past, present, and future. They can be illustrated by the following examples.

- (3) He is working (progressive aspect expressing present time).
- (4) He was working (progressive aspect expressing past time).

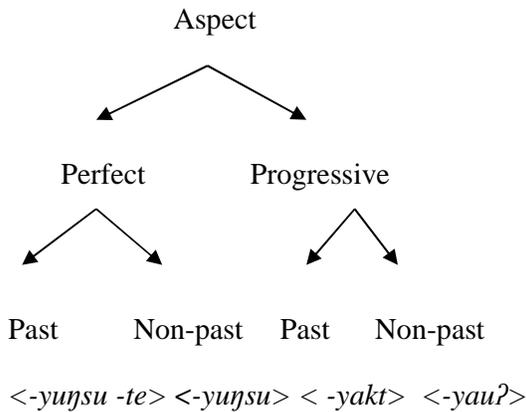
- (5) He will be working (progressive aspect expressing future time).
- (6) He has worked/he has broken a cup (perfect aspect expressing present time).
- (7) He had worked/he had broken a cup (perfect aspect expressing past time).
- (8) He will have worked/he will have broken a cup (perfect aspect expressing future time).
- (9) He has been working (perfect progressive aspect expressing present time).
- (10) He had been working (perfect progressive aspect expressing past time).
- (11) He will have been working (perfect progressive aspect expressing future time).

In English, there are nine subtypes of aspect which play an important role in English grammar and are sometimes more problematic for EFL students.

2.2. Mugali Rai Aspect

Mugali Rai [lmh] is one of the Kirati languages spoken in the small village of Dhankuta District, East Nepal. It is an SOV language with postposition, no gender, verbal affixation markers for person, number, tense, and aspect (Epele, 2012, p. 60), but it has not been recorded by the latest census report (2011). So, it is on the verge of extinction.

In Mugali Rai, there are two types of aspect, perfect and progressive. Both of these can be categorised into two subtypes on the basis of time dimensions, viz. past and present only. There is no grammatical marker referring to future in Mugali Rai, as described below (Rai, 2012, p. 61-69; Rai, 2015, p. 92-95).



2.2.1. Perfect aspect

Perfect aspect can be categorised into two types on the basis of time dimension, which are past perfect and present perfect.

2.2.1.1. Past perfect

In Mugali Rai, the past perfect is constructed through a complex structure. For past perfect, the present perfect marker <-yuŋsu>, and past marker <-ŋe/yē> and another marker <-te> are suffixed to the verb stem gradually. It can be shown in the following examples.

- (12) *kaŋa cama thukyūŋsuyēte*

ka-ŋa cama thuk-yuŋsu-yẽ-te
 ISGA/S-ERG rice cook-PERF-PST-PERF
 I had cooked rice.

- (13) *ka imyuŋsuyẽte*
 ka im-yuŋsu-yẽ-te
 ISGA/S sleep-PERF-PST-PERF
 I had slept.

2.2.1.2. Present perfect

Similarly, the present perfect aspect is expressed by <-yuŋsu> morpheme which is suffixed to the verb stem. The morpheme <-yuŋsu> is immediately followed by the past tense marker <-yẽ>. It can be illustrated by the following examples.

- (14) *kaŋa cama thukyũsuyẽ*
 ka-ŋa cama thuk-yuŋsu-yẽ
 ISGA/S-ERG rice cook-PERF-PST
 I have cooked rice.

- (15) *ka imyuŋsuyẽ*
 ka im-yuŋsu-yẽ
 ISGA/S sleep-PERF-PST
 I have slept.

2.2.2. Progressive aspect

Progressive aspect refers to an ongoing action of verb in Mugali Rai which can be categorised into two types on the basis of time dimension, i.e. past progressive aspect and present progressive aspect, as described below.

2.2.2.1. Past progressive aspect

In Mugali Rai, the morpheme <-yakt> refers to past progressive aspect which attaches to the verb stem immediately, and it is followed by tense and other markers. It is illustrated with the following examples.

- (16) *ka cama thukyaktayẽ*
 ka cama thuk-yakt-yẽ
 ISGA/S rice cook-PROG-PST
 I was cooking rice.

- (17) *ka imyaktaye*
 ka im-yakt-ye
 ISGA/S sleep-PROG-PST
 I was sleeping.

2.2.2.2. Present progressive aspect

In Mugali Rai, the suffix <-yau?>, which immediately follows a verb, represents the present progressive aspect. The following examples make them clear.

- (18) *ka cama thukyau?ŋa*
 ka cama thuk-yau?-ŋa
 1SGA/S rice cook-PROG-1SNPST
 I am cooking rice.
- (19) *ka imyau?ŋa*
 ka im-yau?-ŋa
 1SGGA/S sleep-PROG-1SGNPST
 I am sleeping.

3. A Contrastive Analysis and Acquisitional Problems

English and Mugali Rai are completely different languages. English as an SVO language is a member of the Germanic group of Proto-Indo-European language family (Yule, 1993, p. 168), whereas Mugali Rai as an SOV language is a member of Eastern Kiranti of Tibeto Burman language family (Winter, 1991, p. 110). These two languages are completely different from each other from the perspective of language family. So, the construction of the aspect in both languages is different. If there are differences between them, there would be some acquisitional problems that Mugali Rai as EFL learners face. The major contrasts in constructing aspect between the two languages are as follows.

In the Mugali Rai language, many things such as tense, aspect, person, number, agent, and patient markers are inflected in a single verb, which is a basic feature of the Tibeto-Burman languages. Such features are found in pronominalized languages. It is also known as a *polysynthetic* language (Crystal, 2008, p. 374). Example (20) shows the construction of past perfect in Mugali. Example (21) shows the construction of present perfect in Mugali. Example (22) shows the construction of past progressive in Mugali. Example (23) shows the construction of present progressive in Mugali. There is no future aspect in Mugali Rai like English. Similarly, there is no construction of perfect progressive combination in Mugali Rai, which is found in English. The following examples make them clear.

- (20) *kaŋa cama thukyũsuyẽte*
 ka-ŋa cama thuk-yũsu-yẽ-te
 1SGA/S-ERG rice cook-PERF-PST-PERF
 I had cooked rice.
- (21) *kaŋa cama thukyũsuyẽ*
 ka-ŋa cama thuk-yũsu-yẽ
 1SGA/S-ERG rice cook-PERF-PST
 I have cooked rice.
- (22) *ka cama thukyaktayẽ*
 ka cama thuk-yakt-yẽ
 1SGA/S rice cook-PROG-PST
 I was cooking rice.
- (23) *ka cama thukyau?ŋa*
 ka cama thuk-yau?-ŋa
 1SGA/S rice cook-PROG-1SNPST

I am cooking rice.

In Mugali Rai, markers for all types of aspect follow a stem. The stem *thuk-* is followed by morphemes <-yunsu-PST-*te*> for past perfect, by <-yunsu> for present perfect, by <-yakt> for past progressive, and by <-yau?> for present progressive. In this way, there are only two aspects in Mugali Rai, namely, perfect and progressive. Both perfect and progressive can be categorised into present and past in terms of time dimension. All these aspects in Mugali Rai are morphologically marked.

On the contrary, aspect is not only morphologically marked in English. There are complex constructions for aspect in English. There are three types of aspect, namely, perfect, progressive, and perfect progressive aspect. In English, aspect is constructed through the combination of the auxiliary verb and the forms of the main verb. The structure of aspect in English is more complex in comparison with the structure of Mugali Rai aspect. In English, perfect aspect is constructed through '*have+past participle*'. But there are three types of perfect in terms of time dimensions, which are past perfect, present perfect, and future perfect. Past perfect is constructed through '*had+past participle*', present perfect is constructed through '*have/has+past participle*', and future perfect is constructed through '*shall/will have+past participle*'. Similarly, progressive is constructed through '*be+present participle*'. Progressive aspect is also categorised into three types in terms of time dimensions: past progressive, present progressive, and future progressive. Past progressive is constructed through '*was/were+present participle*', present progressive through '*am/is/are+present participle*', and future progressive is constructed through '*shall/will be+present participle*'. Example (24) refers to past perfect, the examples (25-26) refer to present perfect, and the examples (27-28) refer to future perfect. Similarly, the examples (29-30) refer to past progressive, the examples (31-33) refer to present progressive, and the examples (34-35) refer to future progressive. Similarly, Example (36) refers to past perfect progressive, examples (37-38) refer to present perfect progressive, and examples (39-40) refer to future perfect progressive. The following examples make it clear.

(24) I/we/you/he/she/(it)/they had cooked rice.

(25) I/we/you/they have cooked rice.

(26) He/she/(it) has cooked rice.

(27) I/we shall have cooked rice.

(28) You/he/she/(it)/they will have cooked rice.

(29) I/he/she/(it) was cooking rice.

(30) We/you/they were cooking rice.

(31) I am cooking rice.

(32) We/you/they are cooking rice.

(33) He/she/(it) is cooking rice.

(34) I/we shall be cooking rice.

- (35) You/he/she/(it)/they will be cooking rice.
 (36) I/we/you/he/she/(it)/they had been cooking rice.
 (37) I/we/you/they have been cooking rice.
 (38) He/she/(it) has been cooking rice.
 (39) I/we shall have been cooking rice.
 (40) You/he/she/(it) will have been cooking rice.

In this way, English aspect is more complicated in comparison with aspect in Mugali Rai. Example (24) is not so problematic for Mugali Rai because its equivalent is found in Mugali Rai. It is assumed that structural differences are the source of difficulty in foreign language learning. Lado (1957) viewed learning difficulty and differences as being directly and proportionally related. Of the L2 learner he wrote: "Those elements that are similar to his native language will be simple for him and those elements that are different will be difficult" (as cited in James, 1980, p. 188). So, structural differences between two languages play an important role in learning a foreign language. Structural differences can be found in many examples in English which are not found in Mugali Rai. For instance, the sentences found in examples (25-26) are problematic since there are two auxiliary verbs *have* and *has* which are selected on the basis of subject. Both sentences are realized by a single sentence in Mugali Rai. So, Mugali Rai as EFL learners confuse the auxiliary verbs *has* and *have*.

Similarly, the sentences found in examples (27-28) are not found in Mugali Rai. There is no separate structural form to express future perfect in Mugali Rai. So, there is acquisitional problem for Mugali Rai as EFL learners.

The sentences found in examples (29-30) are also problematic for Mugali Rai as EFL learners. They cannot use *was* and *were* appropriately. The sentences found in examples (31-33) are also problematic in the sense that Mugali Rai as EFL learners cannot use *am/is/are* appropriately.

There is no future progressive in Mugali Rai. So, the sentences found in examples (34-35) are also problematic for them. Mugali Rai learners cannot use *shall* and *will* appropriately.

The sentences found in examples (36-40) are also problematic because there is no combination of both perfect and progressive aspect in the Mugali Rai language.

In conclusion, there are only four structures to express aspect in Mugali Rai, which include past perfect, present perfect, past progressive and present progressive. On the contrary, English has 17 different structures to express aspect. They have been described in examples from 24 to 40. There are considerable structural differences between English and Mugali Rai aspect.

4. Conclusions

There are considerable differences between English and Mugali Rai aspect. So, it is assumed that Mugali Rai learners find English aspect difficult to use appropriately. They can

commit errors in the use of English aspect. So, language teachers should pay a special attention to those areas where there are structural differences between these two languages.

For addressing grammatical errors, Cowan (2009, pp. 45-46) provides extensively detailed guidelines. The question is what we should do about the grammatical errors made by students. To answer the question, we must look at the development of interlanguage. Interlanguage (IL) refers to the language system that evolves as a learner studies an L2. At any given point in the development of a learner's IL, some aspects of the IL grammar may be identical to L2 grammar, but other aspects will be different. The eventual result of instruction and years of practice of speaking an L2 will produce an IL we can call the *end-state grammar*. This is the grammar that the learners will use for communicating with native speakers of the L2 from that point on. It will not change much, and it probably will not be identical to the grammar of a native speaker, but it may be very close to it. Thus, according to this model, the process of learning English is the process by which a learner begins to develop an IL, and that IL continues to grow more and more similar to the English grammar of a native speaker until it stabilizes.

Speaker A: L1.....IL1.....L2 (English)

Speaker B: L2.....IL2.....L2 (English)

Speaker C: L3.....IL3...L2 (English)

This diagramme illustrates the fact that ILs of different learners vary in their approximation to the grammar of the L2 that is being learned.

Some common errors made by Mugali Rai as EFL learners which were taken during my fieldwork in 2011 are as follows:

41. *It is/was looking good.
42. *I am/was hopping.
43. *I am/was understanding.
44. *I am/was resembling.
45. *I lived here since 2000.
46. *He have eaten rice.
47. *You was writing a letter.

Students' proficiency in English aspect can be developed in English through different activities. Mugali Rai as EFL learners cannot make distinction between activity verbs and stative verbs. Stative verbs cannot express progressive aspect. For present progressive, students are asked to describe the present situation of weather like *it's raining*, *the temperature is increasing*, etc. For the correct use of the auxiliary verb in present progressive, students can be categorised into two groups 'A' and 'B'. Under group 'A', students should be categorised into '*I, we, you, he, she, it, they*' groups. On the other hand, under 'B' group, students should be categorized into '*am, is, are*' groups. '*T*' subgroup from group 'A' matches '*am*' subgroup of group 'B'. In this case, when students from 'A' group say '*T*', students from group 'B' say '*am*'. Similarly, '*we*', '*you*', '*they*' subgroups of group 'A' matches '*are*' subgroup of 'B'. When students from group 'A' say '*we*', '*you*' and '*they*', students from group 'B' say '*are*'. Similarly, '*he*', '*she*', '*it*' subgroups of group 'A' match the '*is*' subgroup of 'B'. When

students from group 'A' say 'he', 'she', 'it', students from group B say 'is'. Then, they should be asked to drill until they remember these rules perfectly.

Similarly, the students are asked to perform interview such as, *Where was your father working? What were you doing when the bomb exploded?* This type of interview helps to use past progressive.

Similarly, students are asked to make distinction between sentences like *I have lived here since 2005.* vs. **I lived here since 2005.* Why is the former sentence correct and not the latter one? In this way, students can be engaged in different activities such as dialogue, interview, role play, demonstration, etc. addressing the English aspects system.

Abbreviations

I = first	PERF = perfect
A = agent	PROG = progressive
ERG = ergative	PST = past
NPST = non-past	S= subject
SG = singular	P = patient

References

- Cowan, R. (2009). *The Teacher's Grammar of English: A Course Book and Reference Guide*. India: Cambridge University Press.
- Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics*. The UK: Blackwell publishers.
- Eppele, J., Lewis, M., Regmi, D., & Yadava, Y. (2012). *Ethnologue: Languages of Nepal*. Kathmandu: Linguistic survey of Nepal and Tribhuvan University.
- Given, T. (2001). *Syntax: An Introduction*. Vol I.. Amsterdam: John Benjamins publishing company.
- Givón, T. (2001). *Syntax: an introduction* (Vol. 1). John Benjamins Publishing.
- James, C. (1980). *Contrastive Analysis*. Harlow: Longman.
- Lado, R. (1957). *Linguistics Across Cultures*. Ann Arbor: University of Michigan.
- Payne, J. L., Hand, M., Barovich, K. M., & Wade, B. P. (2008). Temporal constraints on the timing of high-grade metamorphism in the northern Gawler Craton: implications for assembly of the Australian Proterozoic. *Australian Journal of Earth Sciences*, 55(5), 623-640.
- Rai, I. P. (2007). Personal Pronouns in English and Chhiling: A Contrastive Study in *Journal of NELTA*, Vol. 12, No. 1-2, 102-107
- Rai, I. P. (2011). *A Sketch Grammar of Mugali Rai*. A Report Submitted to NFDIN, Lalitpur.
- Rai, I. P. (2012). English Tense and Aspect: EFL Problems that Chhiling Learners Have. *Journal of NELTA*, Vol. 17, No. 1-2, 61-69.
- Rai, I. P. (2015). Mugali and English Tenses: EFL Problems and Pedagogical Strategies. *Nepalese Linguistics*, Vol. 30, 92-95.
- Winter, W. (1991). *The Rai of Eastern Nepal: Ethnic and Linguistic Grouping. Finding of the Linguistics Survey of Nepal and CNAS*. Kathmandu: CNAS.
- Yule, G. (1993). *The Study of Language*. Great Britain: Cambridge University Press.

The Two and a Half Faces of Bangla Monosyllables

Mina Dan¹

mina.dan@gmail.com

¹Department of Linguistics, University of Calcutta, India

Abstract

Monosyllables are syllables as well as words. As syllables, they have a phonetic and a phonological representation - both together facilitate, first, checking legal combinations of segments in a language by devising a template, and second, revealing the phonotactics of the language. As words, monosyllables display at least four distinct features – meaning, overt structure, naming potentiality and orthography. These features assign monosyllables the status of an interface unit too. The present paper attempts to capture the two faces, viz., the syllable-like face and the word-like face of Bangla/Bengali monosyllables on various levels of language structure and treats the interface identity of them as an additional half face of the unit, as mentioned in the title of the paper.

Keywords: monosyllables, Bangla, rime, template, metrical phonology

1 Introduction

On the basis of their phonological component, more specifically the syllabic component, words are categorised as monosyllabic and polysyllabic, the latter being a cover term for disyllabic, trisyllabic, quadrisyllabic and so on. The self-explanatory attribute ‘monosyllabic’ designates the words that are composed of only one syllable. For example, words like [pa] ‘foot’, [am] ‘mango’, [o] ‘that one’, [gan] ‘song’, [ʃon] ‘you-int. listen’, [gram] ‘village’, etc., in Bangla. Until recently, the very concept of monosyllable was understood as a tool for phonological measurement, both quantitative and qualitative. Quantitatively, it marked the words, the real units of speech, with a minimum number of syllable-count, and qualitatively it identified the possible and impossible concatenation of sounds in syllables in a language. However, the uniqueness of monosyllabic words as units with a range of overlapping dimensions was explicitly unfolded by the researchers in the conference titled *Monosyllable: From Phonology to Typology*, organised by and held at Universität Bremen, Germany, in September 2009. With this backdrop, in the present paper we shall point to the uniqueness of monosyllables in Section 1, capture the various dimensions of Bangla/Bengali monosyllables in Section 2, draw a few generalizations on the segmental makeup of Bangla monosyllables in Section 3, and present a brief typological observation on Bangla monosyllables in the concluding section.

1.1 The Uniqueness of Monosyllables

According to the popular definition, syllables are units of pronunciation typically larger than a single sound and smaller than a word, whereas monosyllables, though a type of syllable, are never less than a word and they are often equal to a single sound. Thus, monosyllables are, universally, not only a kind of syllable but also a type of word.

As a kind of syllable, they are recognised as the lowest level macro-units of phonology, the micro-units being the sounds/phonemes. Thus, they have a phonetic as well as a phonological representation. Both the representations together put forward a field-test for, firstly, checking legal combinations of segments in a language, and secondly, framing a set of

laws for pointing at illegal partners in that same linguistic system. For example, in Bangla, as regards legal combinations, a syllable can begin with at the most three consecutive consonants, as in the word [stri] ‘wife’ but never with more than that. The law expressing this is *#CCCC.

In phonology, the former is achieved by devising a template that operates as a syllable-detector while the latter by detecting relevant constraints or template conditions which reveal the phonotactics¹ of the language. In the phonological literature, the framing of tools like syllable template and template conditions banks heavily on the various aspects of the concept of sonority.

Monosyllables are also identified as a type of word as they display at least four distinct features that are unlike syllable-features and like word-features. We would like to interpret the attribute *mono* (to *syllable*) in the term *monosyllable* as an acronym MONO that represents the four features as follows. M stands for meaning, O for overt structure, N for naming potentiality and O for orthography. Hence monosyllables are syllables which have meaning, overt structure, naming potentiality and orthographic implication.

First, unlike syllables and like morphemes, monosyllables have meaning – a feature that makes them morphologically relevant. For example, [din] ‘day’, [lal] ‘red’, [ke] ‘who’, [æk] ‘one’, etc. are morphemes as well as monosyllables with distinct meaning in Bangla.

Secondly, unlike syllables, monosyllables display overt internal structures at different levels – the levels of phonology and morphology. For example, the word [ʃi:p] ‘sheep’ in the English sentence *The sheep are grazing* is a morphologically complex monosyllable as it is composed of two morphemes, sheep and plural.

Thirdly, unlike syllables and like words, monosyllables can name something or some person, i.e., they are words with all privileges of a dictionary entry. One does indeed find all the monosyllables of one’s language in a dictionary of that language.

Fourthly, unlike syllables, each monosyllable is directly linked to a particular reference point in the orthographic convention of a language that has a script and lacks spelling pronunciation. For example, in English, the syllable [si:] has no fixed orthographic reference; it may refer to <sea> as in the word ‘season’, to <cea> as in ‘ceasing’, to <see> as in ‘seeded’, to <sei> as in ‘seize’, to <ce> as in ‘cedar’, or to <cei> as in ‘ceiling’. But the monosyllable [si:] is linked either to <sea> when it is a noun, or to <see> when it is a verb.

These four features assign monosyllables the status of an interface unit too. The features mentioned above, viz., the syllable-like and the word-like features, are the general features of monosyllables in many languages. Monosyllables also present language-specific features. The present paper attempts to capture the language-specific features of Bangla monosyllables.

1.2 Bangla Monosyllables

The literature on Bangla metrical analysis and Bangla linguistics presents a few studies that deal with various aspects of Bangla syllables. Prabodh Chandra Sen, the grammarian metricist researching since 1922, identified syllables as the key units determining

various metrical patterns in Bangla and illustrated meticulously the basic concepts such as syllable, quantity-unit, mora, etc. (Sen, 1986). However, his studies hardly report any observation on Bangla monosyllables.

In the field of Bangla linguistics, Mallik (1960) and Hai (1964) deal with Bangla consonant clusters and sequences which have some bearing on the syllabification rules in Bangla. Hai (1964) also presents syllable-wise time duration of a handful of Bangla words using kymograph tracing. Sarkar (1979, 1986) throw light on the structure of Bangla syllables. Sarkar (1985, 86) illustrates Bangla diphthongs which have some relevance in the rime structure of syllables. Mallik et al. (1998) provide a frequency count of various syllables at different positions in words, and Dan (2004) presents a characterization of Bangla syllables based on the theory of metrical phonology. However, none of these works concentrates on Bangla monosyllables.

The current section aims to highlight the language-specific features of Bangla monosyllables with least theoretical complexities. It attempts to capture the two faces, the syllable-like face and the word-like face, of Bangla monosyllables by marking the features of monosyllables on various structural levels of the language, namely, phonetic, phonological, interface, morphological, syntactic, and lexical. The interface identity of Bangla monosyllables may be treated as an additional half face of the unit, as mentioned in the title of the paper. While the phonetic and phonological levels are directly related to the syllable-like face, the morphological, lexical, and syntactic levels are related to the word-like face of monosyllables.

2.1 Phonetic

Phonetic features include stress, vowel length, and aspiration.

2.1.1 Stress

Linguists on impressionistic grounds assessed Bangla as a syllable-timed language like Spanish and French and unlike German and English that are stress-timed. Stress is not phonemic in Bangla. According to Chatterji (1928) and Bykova (1981), stress plays no special role in Bangla so far as the individual words are concerned. Hayes and Lahiri (1991) state that stress in Bangla is predictable and dominantly initial in isolated words and is always subsidiary to sentence-stress.

2.1.2 Vowel length

Bangla vowels lack phonemic length. However, they have purely predictable phonetic length. In monosyllables, vowels tend to be long while in polysyllabic words non-long. For example, [du:d] ‘milk’, [b^ha:t] ‘rice’, but [dudb^hat] ‘milk and rice’, [t^hi:k] ‘right’, but [t^hikt^hak] ‘okay’, [ti:n] ‘three’, but [tintɛ] ‘three-Cl’, etc. Thus, length in Bangla monosyllables adds to the phonetic duration of monosyllables, the arithmetical aggregate of the duration of the constituent segments.

2.1.3 Aspiration

Aspiration is phonologically relevant in Bangla, it is phonemic. All stop and affricate consonants have their aspirated counterparts, e.g. [k-k^h, g-g^h, t-t^h, d-d^h, p-p^h, b-b^h, c-

c^h, j-j^h]. However, word finally aspirated segments tend to lose aspiration almost entirely and appear to be unaspirated. For example, [du:d] ‘milk’, but [dud^her] ‘of milk’, [mat] ‘field’, but [mat^he] ‘in the field’, etc. Compared to Hindi, a language in which monosyllables retain perceptually very strong aspiration word finally, e.g., [ʃud^h] ‘pure’, [dugd^h] ‘milk’, etc., Bangla monosyllables present a very mild, almost negligible amount of aspiration in their coda positions.

2.2 Phonological

Phonological features concern, on the one hand, the concept of markedness and on the other, the system of mora count in the domain of poetry.

2.2.1 *Less marked vs. more marked properties*

A structure that is avoided in unrelated languages is said to be marked relative to those structures that are not avoided. The structures of Bangla monosyllables present the following less marked properties:

- i) Simple onsets are allowed, e.g., *ke* ‘who’.
- ii) Monosyllables may be open, e.g., *ke* ‘who’, *na* ‘not’.
- iii) The most preferred syllable pattern is CV.
- iv) Monophthongs serve as peaks, e.g., *kan* ‘ear’.
- v) Complex codas are not allowed in native monosyllables.
- vi) They conform to the Sonority Sequencing Principle.²

More marked properties are:

- i) Onsetless syllables are allowed, e.g., *aj* ‘today’.
- ii) Complex onsets are allowed, e.g., *b^hru* ‘eyebrow’.
- iii) Diphthongs may serve as peaks, e.g., *teif* ‘23’.
- iv) Closed syllables are allowed, e.g., *mon* ‘mind’.
- v) Simple codas are allowed, e.g., *g^hum* ‘sleep’.
- vi) In borrowed monosyllables complex codas are allowed, e.g., *bæŋk* ‘bank’.
- vii) s+stop, the cross-linguistically attested cluster that violates the Sonority Sequencing Principle, is available, e.g., *st^hir* ‘motionless’.

2.2.2 *Controversy of more marked peaks*

Monosyllables may have any of the (i) seven oral and (ii) seven nasal vowels and (iii) diphthongs as their peak. Among these three sets of segments, diphthongs are the more marked as well as controversial peaks.

Often a word with a diphthong as the peak followed by a simple consonant is ambiguous between two readings, monosyllabic and disyllabic, as the vocalic slot reflects the phonetic length, the usual phenomenon within a monosyllable, e.g., *deul* and *de-ul* ‘temple’, *baif* and *ba-if* ‘22’, *gour* and *go-ur* ‘a male name’.

Contrary to this, a diphthong followed by a single consonant followed by another vowel (in a disyllabic word, of course) is shorter and maintains its diphthongal character, e.g., *deu-le* ‘insolvent’, *bai-fe* ‘on 22nd’, *gou-ri* ‘a female name’.

The script convention too contributes to this controversy. Bangla script has unitary graphemes for only two diphthongs, viz. [oɪ] and [ou]. Hence, while these two are visibly single units others are not, e.g. *pouf* ‘name of a Bengali month’ <xx> but *k^heur* ‘vulgar song’ <xxx>.

In fact, *k^heur* is more prone to a disyllabic reading.

Another interesting gap in the system is that Bangla does not have any monosyllable with [oɪ] as its peak followed by a single consonant, i.e. *(C)oɪC. So the scope of verifying the influence of script is restricted too.

However, double reading does not assign a questionable status to Bangla diphthongs altogether, because the monosyllables with word final diphthongal peaks categorically yield single reading regardless of the difference in their visual representations, e.g. *bou* ‘wife’ <x>, *doi* ‘curd’ <x>, *b^hoḡ* ‘fear’ <xx>, *mæo* ‘baby-talk item for cat’ <xx>, *hae* ‘alas’, etc.

The interesting point here is that we do not really know whether or not *dour* and *k^heur*, with reference to syllabification, behave the same way. As regards this, even speakers’ opinions are not of much help – as out of the following four options only the last one would be rejected outrightly, whereas the other three would receive mixed responses.

Between *dour* and *k^heur*

1. Both are monosyllabic.
2. Both are disyllabic.
3. *dour* is monosyllabic, while *k^heur* is disyllabic.
- *4. *k^heur* is monosyllabic, while *dour* is disyllabic.

The above issue entails a valid question: do we really know the actual range of monosyllables in Bangla? Researches in this direction are yet to be taken up.

2.2.3 *Indigenous metrical literature*

In terms of mora counting, monosyllables receive the same treatment as any other syllable in the indigenous metrical literature. Bangla has three basic metrical patterns, viz. syllabic, moric and composite. A syllable with a final simple vowel, conventionally known as an open or light syllable, is considered monomoric in all these three patterns. But a syllable with either a coda or a diphthongal peak, known as a closed or heavy syllable, shows different mora count in different metrical patterns – in the syllabic pattern it is monomoric, in the moric pattern it is bimoric, and in the composite pattern it is positionally bimoric (word-finally) or monomoric (elsewhere). Thus, the closed or heavy syllables are of greater interest than the open or light ones in metrical literature.

2.3 Interface Properties

These properties function in the interface area between phonology and morphology and include a few phonological processes, viz. vowel harmony, deletion and semivowel formation, which may operate even within the domain of monosyllables.

2.3.1 Vowel harmony

In this phonological process, a high vowel trigger, viz. *i* and *u*, raises the vowel height of the syllable immediately preceding it by one degree. For example, *ʃo* ‘sleep’ + *i* ‘1Pres’ = *ʃui* ‘I sleep’.

Vowel harmony is so pervasive in Bangla that the process even blocks [**æ̃*, **æ̃̄*, **ɔ̃*, **ɔ̃̄*] diphthongs as peaks of monosyllables.

2.3.2 Deletion

A suffixal vowel is deleted in morphologically complex monosyllables, e.g. *k^ha* ‘eat’ + *en* ‘3Pres-For’ = *k^han* ‘s/he eats’.

The process of deletion has other morphological implications too. Deletion applies to the verbal sector and invariably converts CV-VC into monosyllabic CVC (*k^ha-en* = *k^han*). As a consequence of this, the canonical shapes CVVC (e.g. *teif* ‘23’) and VVC (e.g. *ain* ‘law’) of monosyllables are rightly predicted as belonging to the nominal sector of the language.

2.3.3 Semivowel formation

In semivowel formation, a simple peak becomes complex, i.e. diphthongal, e.g. *ho* ‘be’ + *o* ‘Pres-Ord-Imp’ = *hoɔ* ‘you become’.

2.4 Morphological

Morphological properties refer to the morphological structure and grammatical category of monosyllables.

2.4.1 Structurally simple or complex

Morphologically, monosyllables may be simple or complex. Examples of simple forms are *nak* ‘nose’, *boi* ‘book’, *maf* ‘month’, *rɔŋ* ‘colour’, *nil* ‘blue’, *k^hub* ‘very’, *stri* ‘wife’, *st^han* ‘place’, etc.

Morphologically complex monosyllables are almost always of the portmanteau type, where the number of morphemes exceeds the number of morphs. For example, *k^ha* ‘eat’ + *o* ‘2Pres-Ord-Imp’ = *k^haɔ* ‘you eat’, the monosyllable is composed of two morphs but five morphemes.

2.4.2 Part of speech

Monosyllables in Bangla may belong to any part of speech, e.g. noun (*din* ‘day’, *rat* ‘night’, *gac^h* ‘tree’, etc.), pronoun (*ʃe* ‘s/he-Ord’, *tui* ‘you-Inf’), verb (*naɔ* ‘take’, *bɔl* ‘speak’), adjective (*t^hik* ‘right’, *bɔd* ‘bad’, numerals from 1 to 10, etc.), adverb (*k^hub* ‘very’, etc.), postposition (*pɔr* ‘after’, etc.), and others (*ba* ‘or’, *na* ‘not’, *hã* ‘yes’, *c^hih* ‘fie!’, *to* ‘so’, etc.).

2.5 Syntactic

Syntactic implications include the case assignment and verbal agreement properties and the holophrastic aspect of Bangla monosyllables.

2.5.1 Case assignment

Monosyllables may assign case to a relevant syntactic category, e.g. the monosyllabic postposition *pər* ‘after’ assigns genitive suffix to its immediately preceding noun, *gan-er pər* ‘song-Gen after’ ‘after the song’.

2.5.2 Verbal agreement

In Bangla, monosyllables may play an active role in verbal agreement. The verbal agreement may vary depending on a monosyllabic subject of the sentence, e.g. *tuj lik^hbi* ‘you-Inf write-will’ ‘you will write’, but *je lik^hbe* ‘s/he-Ord write-will’ ‘s/he will write’.

2.5.3 Holophrase

A holophrase is a one-word sentence. In Bangla, monosyllables may be used as holophrases, e.g. *ma* ‘Oh mother!’, *ki* ‘What?’, *to* ‘So what?’, etc.

2.6 Lexical

Bangla has a segmented lexis. With reference to the source, the words, including the monosyllables, are commonly classified into three strata:

- 1) Direct adoptions from Sanskrit, traditionally termed *tatsama* ‘equal to that’ (*tat* ‘that’, *sama* ‘equal’). Examples of monosyllables are, *rup* ‘beauty’, *gun* ‘quality’, etc.
- 2) Native Bangla words rooted in Sanskrit and Prakrit, termed *tadbhava* ‘derived from that’ (*bhava* ‘derived’), e.g. *duj* ‘two’, *tin* ‘three’, *car* ‘four’, *pāc* ‘five’, etc.
- 3) Other borrowed stocks that include Turko-Perso-Arabic items and items from English, Dutch, Portuguese, French, Chinese, Gujarati, Tamil and other, even forgotten, origins, e.g. *lajn* ‘line’, *dost* ‘friend’, *ca* ‘tea’, *ṭæks* ‘tax’, etc.

Bangla has a very receptive nature as far as the fortification of lexis is concerned. These three strata build up the active vocabulary of the native speakers. In the borrowed stock, degrees of naturalization are identified, e.g. *lajn* ‘line’, *bæŋk* ‘bank’ do not stand out as loans, while *jerəks* ‘xerox’ shows semi-naturalization, and *səffoær* ‘software’ almost zero-naturalization (Dasgupta, 2003). While commenting on lexical diversity of Bangla, Dasgupta (2003) uses the numerals, a representative sector, and notes that in Bangla

- i) cardinals are *tatsama* at 0 and 1
- ii) *tadbhava* from 2 to 100
- iii) Persian at 1000
- iv) *tatsama* for ordinals from 1 to 10
- v) *tadbhava* for date (*pəɽla* ‘1st’, *dofra* ‘2nd’ ...) and playing cards cardinals (*duri* ‘2’, *tiri* ‘3’ ...), and

- vi) English for ordinals in academic and other competitive contexts (1st, 2nd, 3rd)

Bangla phonology comprises reflections of all these stocks, e.g. complex onsets identify loans from Sanskrit, English, etc. while complex codas loans from English and Persian. Thus, the canonical shape of a monosyllable gives clue for the strata which it belongs to.

3 Generalizations

The current section will present a few general observations on the canonical shapes and optimal pattern of Bangla monosyllables. Though in this regard the tools and concepts of the theory of metrical phonology play a significant role, we decide to stay away from the theoretical complications as far as practicable in this paper. The data and claims made in this paper are empirical in the sense that anyone knowing Bangla as one's 1st or 2nd tongue can always check them intuitively – a fact that well justifies our decision.

3.1 Canonical Shapes

Though a number of studies have already highlighted the various aspects of Bangla syllables, for example, the phonetic length and sequences at the boundaries of syllables (Hai, 1964), the segmental composition and frequency of occurrence of syllables (Mallik, 1960; Mallik et al., 1998), the canonical shapes and their metrical relevance (Sarkar, 1979, 1986), analysing Bangla syllables in the metrical framework (Dan, 2004) and in the Optimality framework (Kar, 2009), none of the studies provides any statement on monosyllables.

The studies mentioned above attest 16 canonical shapes of Bangla syllables of which 15 are found in monosyllables. The canonical shapes may be broadly categorised into three groups, viz. vowel final, e.g. *ne* 'you-Inf take', diphthong final, e.g. *b^hai* 'brother', and consonant final, e.g. *p^hul* 'flower'. The last group may further be categorised into two sub-groups, viz. with a simple vowel as the peak, e.g. *p^hul* 'flower', and with a diphthong as the peak, e.g. *teif* '23'. A notable point here is that the initial consonant or consonants, i.e. the onset part, hardly play any role in this categorization, a point that gained sufficient justification in the fields of both phonology and metrical study.

In the phonological literature, the diphthongal peaks of Bangla monosyllables are represented as VV, more visibly V₁V₂, sequences, in which

- 1) the V₁ is always [+syll], i.e. a pure vowel
- 2) the V₁ is always higher than the V₂ in term of vowel height
- 3) the V₂ is a semivowel with the feature [-syll], i.e. more like a consonant
- 4) the 2nd V is more like a consonant in many respects, viz. phonetically, morphophonemically, and metrically.

In metrical phonology, the segmental makeup of syllables is represented hierarchically in terms of tree diagrams, where the highest node syllable branches into onset and rime; then at the next level, the rime node branches into peak and coda. Treating every [-syll] segment under the rime as coda, a decision which would not yield any adverse

consequence in Bangla phonology, the tree structures of the above groups and sub-groups of Bangla monosyllables will be as follows:

- i) Vowel final monosyllables will have nonbranching rime, i.e. the rime branch will contain only the peak, the V, but no coda.
- ii) Diphthong final monosyllables will have branching rime, i.e. the rime will branch into the peak, the V₁ and the coda, the V₂.
- iii) Consonant final monosyllables with a diphthongal peak will have branching rime, i.e. the rime will branch into the peak, the V₁, and the coda. Further, the coda will branch into two coda segments, the V₂ and the C.
- iv) Consonant final monosyllables with a simple vowel peak will have branching rime, i.e. the rime will branch into the peak, the V, and the coda, the C. The coda may further branch if it contains more than one consonant.

In metrical phonology, syllables with branching and non-branching rimes are termed heavy and light, respectively. Thus, the diphthong final and both the types of consonant final monosyllables, i.e. (ii), (iii) and (iv), together are considered heavy syllables, while the vowel final ones, i.e. (i) light.

Following the above principle of metrical phonology, viz. syllables with branching rimes are heavy, and those with nonbranching rimes are light. The 15 canonical shapes of Bangla monosyllables may be grouped as light and heavy as follows.

<u>Light</u>			<u>Heavy</u>		
V	e	‘this person’	CVC	gac ^h	‘tree’
CV	ki	‘what’	VC	aṭ	‘8’
CCV	sri	‘Mr.’	CCVC	gram	‘village’
CCCV	stri	‘wife’	CVVC	douṛ	‘run’
			CCVVC	praij	‘prize’
			CVCC	last	‘last’
			CCCVC	strir	‘wife’s’
			VVC	aṛn	‘law’
			VV	aṛ	‘income’
			CVV	boṛ	‘book’
			CCVV	praṛ	‘almost’

3.2 Optimal Pattern

Based on the 15 canonical shapes shown above, a generalization on the structure of Bangla monosyllables may be drawn as follows:

(C)(C)(C)V(V)(C)(C)

It means at the onset position, i.e. before the peak, a monosyllable may have maximum three consecutive consonants and minimum no consonant at all; at the peak position, it may have at the most two segments and at least one, and at the coda position, i.e. after the peak, maximum two consonants and minimum no consonant at all. This may also be expressed as follows:

$$C_0^3V_1^2C_0^2$$

On the basis of the above generalization, we propose below the optimal pattern of Bangla monosyllables:

$$CCCVVCC$$

However, the optimal pattern has no exponent in the language. Other generalities are furnished below.

- i) The most preferred canonical shape in Bangla is CV.
- ii) Complex onsets single out loan monosyllables from English and Sanskrit.
- iii) Complex codas single out loan monosyllables from English and Persian.
- iv) Complex onsets and codas are simplified in the non-standard varieties of Bangla by means of either deletion or epenthesis though they are generally retained in the standard tongue.

4. Conclusion

In conclusion, we provide a few typological statements on Bangla monosyllables in comparison with the syllables of other languages.

- i) Bangla monosyllables like those in Sanskrit and Gujarati register has word-final deaspiration, unlike languages that retain aspiration word finally, such as Hindi.
- ii) Bangla has light and heavy monosyllables, unlike languages with a three-way division of syllables, viz. light, heavy, and extra heavy, as in Arabic.
- iii) Bangla allows both open and closed monosyllables, unlike languages like Shona and Fijian that allow only open syllables.
- iv) Bangla has four semivowels, viz. \dot{i} , u , e , and o , unlike languages with two semivowels, such as English, Hindi, Gujarati, Tamil, Khasi, Angami, etc.
- v) Bangla monosyllables allow restricted voicing in obstruent clusters, both the members being either voiced or voiceless, unlike languages that reject any voicing, such as Maori, Ainu, as well as languages that allow unrestricted voicing, such as Kannada, Tulu, etc.
- vi) In Bangla monosyllables, onsets are optional, unlike languages with obligatory onsets, such as Arabic.
- vii) Bangla monosyllables invariably reject geminate onsets and codas, unlike languages that allow geminate onsets, such as Leti, Trukese, or geminate codas, such as Hungarian.

- viii) Bangla monosyllables allow complex onsets and in borrowed items even complex codas, unlike languages that invariably reject complex onsets and codas, such as Japanese.

Notes

1. Phonotactics determines the set of permissible arrangements of sequences of speech sounds in a given language. It rejects the unacceptable sound sequences by means of phonotactic constraints which are highly language specific. For example, [st] is not a permissible cluster in Japanese, though it is a permissible one in Bangla.
2. The Sonority Sequencing Principle imposes restrictions on segment sequencing in syllables. It states that the peak of a syllable has the highest sonority value, while between the peak and the syllable boundary, on both sides, the sonority values increase as one nears the peak and decrease as one moves away from the peak, e.g. #12321#. Each sound segment is assigned a sonority value on the sonority scale, the low vowels being assigned the highest value while the voiceless stops the lowest. For a detailed discussion, see Dan (2012).

Abbreviations

1	1 st person
2	2 nd person
3	3 rd person
Cl	classifier
For	formal
Gen	genitive
Imp	imperative
Inf	informal
Ord	ordinary
Pres	present tense

References

- Bykova, E. M. (1981). *The Bengali Language*. Moscow: Nauka Publishing House. Central Dept. of Oriental Literature.
- Chatterji, S. K. (1928). *A Bengali Phonetic Reader*. London: University of London Press.
- Dan, M. (2004). *Some Issues in Metrical Phonology of Bangla: The Indigenous Research Tradition*. (e-book). www.ciil-ebooks.net OR <http://www.ciil-ebooks.net/html/mp/content.htm>. Mysore: CIIL.
- Dan, M. (2012). A Sonority Survey of the Monosyllables in Bangla. Otsuka, H., C. Stroh & A. Urdze (eds.) *More Morphologies: Contributions to the Festival of Languages*,

- Bremen, 17 Sep to 7 Oct, 2009. Diversitas Linguarum 35, 67-88. Bochum: Brockmeyer..*
- Dasgupta, P. (2003). Bangla. Cardona, G & D. Jain (eds.) *The Indo- Aryan Languages*, 351-390. London: Routledge,.
- Hai, M. A. (1964). *dhoni biggæn o baᅅla dhoni-totto*. (Phonetics and Bangla phonology). Dacca: Bengali Academy.
- Hayes, B., & A. Lahiri. (1991). Bengali intonational phonology. *Natural Language and Linguistic Theory 9*, 47-96.
- Kar, S. (2009). *The Syllable Srtucture of Bangla in Optimality Theory and its Application to the Analysis of Verbal Inflectional Paradigms in Distributed Morphology*. Unpublished Ph. D. dissertation. Universität Tübingen. Tübingen.
- Mallik, B. P. (1960). Phonemic analysis of the consonant clusters in standard colloquial Bengali. *Bulletin of the Philological Society of Calcutta. 1(2)*, 37-46.
- Mallik, B. P., Bhattacharya, N., Kundu, S.C., & Dan, M. (1998). *The Phonemic and Morphemic Frequencies of the Bengali Language*. Calcutta: The Asiatic Society.
- Sarkar, P. (1979). Syllable-totto o baᅅla b^haᅅfar syllable-er goᅅ^hon. (Theory of syllable and the structure of Bangla syllables). Sengupta (ed.) *bibhab, 10*, (April-June), 33-54.
- Sarkar, P. (1985-86). baᅅla diᅅᅅor d^honi. (Bangla diphthongs). *Bhasha 4-5*, 10-22.
- Sarkar, P. (1986). Aspects of the Bengali syllables. Paper presented in the National Seminar on the Syllable in Phonetics and Phonology. Osmania University. Hyderabad.
- Sen, P. C. (1986). *nuton c^hondo porikkrōma*. (Revised critical survey of metrics). Calcutta: Ananda Publishers.

Use and Usefulness of Monolingual, Bilingual and Semi-bilingual Decoding Dictionaries in Pakistan: A Comparative Study

Rana Faqeer Muhammad Aslam¹, Dr. Zafar Iqbal¹, Naveed Ahmed Chaudhary¹

¹Department of English, Bahauddin, Zakariya University Multan, Pakistan.

Abstract

Theoretical lexicography is concerned with scholarly studies in such disciplines as linguistics, especially Lexicology (Hartman & James, 2002). Research into theoretical lexicography covers dictionary criticism, dictionary history, dictionary needs analysis, dictionary structure, dictionary typology, and dictionary use. The present paper is related to dictionary typology and dictionary use. It presents a comparison of the use and usefulness of three types of the decoding dictionaries used by learners and recommended by teachers of the English language in Pakistan. The methodology applied for the present research is qualitative and descriptive in approach. However, it also presents the results in percentages in order to compensate for the subjectivity of qualitative approach and to reach objective conclusions. The participants of the research include 400 students and 25 teachers belonging to different universities and colleges in Pakistan. Questionnaire as a data collection tool has been exploited for the current research. The comparative results of the three types of decoding dictionaries reveal that the learners mainly utilise two types of dictionaries, i.e. semi-bilingual (L2-L2-L1) English to English and Urdu dictionary with a slight edge over monolingual (L2-L2) English to English dictionary. The research also finds that the learners of English in Pakistan less frequently make use of bilingual (L2-L1) English to Urdu dictionary.

Keywords: theoretical lexicography, dictionary use, decoding dictionary, monolingual, bilingual, semibilingual

1. Introduction

A dictionary is of vital importance for learning and teaching of any foreign language. In Pakistan, English has the status of an official language along with Urdu. It is taught as a foreign language. Different types of decoding and encoding dictionaries are used by the learners. These dictionaries include monolingual, bilingual and semi-bilingual or bilingualised dictionaries.

The paper in hand presents a research conducted in a purely non-native English setting. It focuses on the comparison of the use and usefulness of a monolingual English to English dictionary (henceforth EED), a bilingual English to Urdu dictionary (henceforth EUD), and a semi-bilingual English to English and Urdu dictionary (henceforth EEUD) for the learners of the English language in Pakistan. This study is the result of one of the aspects of the PhD project accomplished by the principal author. The main project was concerned with the use and usefulness of dictionaries in general and semi-bilingual EEUD in particular.

Keeping in view the topic and the nature of its research, a mixed method approach is utilised. Applying survey method, questionnaire is used as a tool for collecting information from four hundred students and twenty-five teachers belonging to different universities and colleges in Pakistan. The analyses of the collected data are presented in the qualitative descriptive form. At the same time, the participants' responses to the closed-ended questions are quantitatively analysed and presented in graphs.

2. Review of Literature

A review of the relevant literature is presented in this part of the paper, which will create a link between the previous and the present studies related to the use and usefulness of dictionaries.

2.1 Types of Dictionaries

Different perspectives are considered to determine the dictionary types. These perspectives are the size, coverage and depth of content, time reference, format and arrangement, functionality, medium of production, and level and skills of the users (Atkins & Rundell, 2008). The paper in hand focuses the types of dictionary related to the coverage of language and functionality.

2.1.1 Types of Dictionaries According to Language Coverage

Types of dictionaries according to language coverage include monolingual, bilingual, multilingual, and semi-bilingual or bilingualised dictionaries. A monolingual dictionary provides all kinds of information about the headwords in the same language. This kind of dictionary is the first source of reference for the native speakers of any language.

A bilingual dictionary covers the vocabulary of two languages. It provides meanings of words in another language. According to Zgusta (1971), the basic purpose of a bilingual dictionary is to coordinate the lexical units of one language with those of another language which are equivalent in their lexical meaning.

A multilingual dictionary presents the vocabularies of more than two languages by giving meanings of all the languages covered. Terminological dictionary is an example of this type of dictionary. It provides vocabulary of various disciplines across languages developed as a result of some international standardization of the technical terms and definitions (Hartmann & James, 2002). A multilingual dictionary is also named as a 'plurilingual dictionary' (Sterkenburg, 2003, p. 409).

A semi-bilingual or bilingualised dictionary 'contains headwords and definitions in one language and translation equivalents in the other language' (Hartmann & James, 2002). It features the qualities of two types of dictionaries, monolingual and bilingual. The terms 'bilingualised dictionary' and 'semi-bilingual dictionary' are alternately used for the same type of dictionary. A semi-bilingual dictionary provides the main information like the headword, part of speech, definition, example of usage in English, and in addition to that, 'each meaning of the headword has a brief translation in the learner's native language' (Kernerman Dictionaries). Semi-bilingual dictionaries have been the subject of investigation in studies like Laufer and Melamed (1994), Nakamoto (1995), Laufer and Hadar (1997), Marelllo (1998), Thumb (2004), Lew (2004), and many more.

2.1.2 Types of Dictionaries According to the Functionality

The purposes of the use of dictionaries determine the classification of dictionaries according to their functionality. The examples include encoding dictionary and decoding dictionary. A decoding dictionary helps perceive a language. It provides the meanings of lexical items a user may need to understand a text. Monolingual dictionaries mainly serve the purpose of a decoding dictionary. A bilingual dictionary also helps in improving the receptive skills of the users. They utilise it in reading and listening contexts, and also in translating a text from foreign language into their own language. Decoding dictionaries are also termed as *passive dictionaries* (Hartmann & James, 2002). On the other hand, an encoding dictionary helps the users with encoding tasks. A thesaurus is an example of encoding dictionary. It gives the lexical options for encoding a message in writing or speaking. A bilingual encoding dictionary fulfils the requirements of the users for translating the native language into the target language. According to Hartmann and James, this type of dictionary is also termed as *active dictionary* (p. 3).

The research in hand is concerned with monolingual, bilingual, and semi-bilingual decoding dictionaries.

2.2 Research into the Use and Usefulness of Dictionaries

The usefulness of dictionaries for the learners has been a subject of study for theoretical lexicographers since the beginning of the research in this area. Barnhart (1962) conducted the first important study in this context. His study was based on the teachers' observation about their students' dictionary-using habits. Although his study is considered a milestone in dictionary research, it did not involve dictionary users themselves. Tomaszczyk (1979) focused actual dictionary users involving students, teachers, and translators. He was the first researcher to examine the dictionary needs of non-native speakers of English. Questionnaire was used by him as a tool for data collection. His study, conducted in Polish and American contexts, found that the bilingual dictionaries were preferred by the participants.

Hartmann (1983) conducted a research into the use of bilingual dictionaries. The participants of his study were 118 learners of German and 17 teachers. He used questionnaire to collect data. He found out that the learners utilised a bilingual dictionary mainly for the purpose of translation. The other contexts of the use of this type of dictionary were reading and writing. He also pointed out that the subjects of his study did not have enough training for using a dictionary.

The differences and the effectiveness of three types of dictionaries, i.e. monolingual, bilingual, and bilingualised were studied by Laufer and Melamed (1994). Their study involved EFL learners' comprehension and production of fifteen low frequency words in English. They found that "the bilingualised dictionary was significantly more effective than the other two" (Laufer & Melamed, 1994, p. 575).

Non-native learners of the English language as users of dictionary were involved as participants in a study conducted by Robert Lew in 2004. A broad sample of 712 Polish learners of the English language and an unspecified number of teachers participated in this study. He

used questionnaire and designed a test to collect data for his study. He utilised monolingual, bilingual, or semi-bilingual dictionaries for the 'Dictionary Effectiveness Test'. He found that the bilingual dictionaries were used more often than the monolingual dictionaries, but the participants gave higher ratings to monolingual dictionaries than the bilingual dictionaries. His study also revealed that the effectiveness of dictionaries increased as the users moved to higher levels. In conclusion to his study, he questioned the validity of the educators' recommendation of the supposed superiority of monolingual dictionary. He was of the view that there was hardly any empirical evidence to support this supposed dogma. According to him, 'if any evidence was available, it pointed to the more effectiveness of the use of bilingual dictionary for reception purposes.' (Lew, 2004, p. 179).

In the backdrop of the studies discussed above, the research in hand has been designed to find out the use and usefulness of decoding dictionaries by the learners of English in Pakistan. It involved students being the direct users of dictionaries, and teachers as the indirect observers. Since the comparative study of the use and usefulness of monolingual EED, bilingual EUD, and the semi-bilingual EEUD in Pakistani context has not gained the attention of the theoretical lexicographer so far, the present study is an attempt to fill this gap.

3. Research Methodology

Mixed method approach has been applied for the present study. It is qualitative descriptive in approach. Describing, recording, analysing and interpreting the existing condition is called a qualitative descriptive research. This type of research deals with the existing trends in a point of time. Singh (2005) believes that qualitative descriptive research uses non quantitative methods and systematic procedures to discover non-quantifiable relationship between existing variables. A qualitative descriptive research has its own limitations due to the subjectivity involved on the part of the researcher. However, this subjectivity can be lessened by the addition of specific results in quantitative form. In order to reach objective conclusions, the study in hand presents the analyses of the responses of the participants to the closed-ended questions in percentages. Subsections to follow provide details of the methodology involved in its concept and framework.

3.1 Research Question

Since the present study is concerned with the comparative use and usefulness of three types of decoding dictionaries, it seeks to find an answer to the question: Which type(s) of dictionary(ies) out of monolingual, bilingual, and semi-bilingual is/are utilised by the learners of the English language in Pakistan?

3.2 Participants of the Research

The participants of this study were the students majoring in English language and literature from different universities and colleges of Pakistan. Four hundred students, consisting of 246 female and 156 male, came from various linguistic backgrounds. They used their local or regional languages in routine communication. English was not the mother tongue of any of the participants, and a large number of the participants did not even communicate in

Urdu for general purposes in their daily lives. However, all of them were competent to use Urdu. It is noteworthy that a number of participants had switched over to Urdu from their mother tongue at school age. The participants had been learning the English language for many years.

Majority of the participants enjoyed the availability of dictionary(ies) at their home and educational institutes and they were conscious of the importance of dictionaries as reference material. They had been using different types of dictionaries for a number of years for learning of the English language.

Teachers, as indirect observers of students using dictionaries, have also contributed to this study with their invaluable opinion. Using questionnaire, the data were collected from twenty-five teachers of the English language belonging to different universities and colleges in Pakistan.

3.3 Data Collection

The main data collection procedure for this study was the survey using the tool of questionnaire, which had been one of the favourite tools among the researchers in lexicography. Although questionnaire has been criticised as a data collection technique in dictionary research by theoretical lexicographers like Hatheral (1984), “it continues to be used as a major tool in any research concerned with people” (Diab, 1990). Considering the widespread use of questionnaire, Robert Lew re-examined the use of questionnaire in dictionary research. He supported its use and asserted that the “other methods should not supplant but rather supplement questionnaires” (Lew, 2002, p. 270). Researchers like Diab (1990), Iqbal (1987), Kipfer (1985), Hartman (1983), Bejoint (1981), Tomaszczyk (1979), Quirk (1973) and Barnhart (1962) have utilised questionnaire as a tool to collect data in studying dictionary use.

Following the trend set by the researchers studying dictionary use, questionnaire as a data collection tool was exploited for the study in hand to find the relative significance of the three types of dictionaries mentioned above. Questionnaire for this study has been designed keeping in view the suggestions put forward by Lew (2002). Considering the limitations and flaws of questionnaire, informal unstructured interviews were also conducted after administering the questionnaire. Iqbal (1987) and Diab (1990) have utilised interview as a tool along with questionnaire for their studies concerned with dictionary use. Along with these, the participant observation, which is an integral part of any kind of qualitative research, has also been helpful.

3.3.1 Questionnaire for Learners of English

Questionnaire used for the main project was designed to draw out information about the learner’s dictionary use in general, owning a dictionary and frequency of using a dictionary, and the helpfulness of information available in a particular kind of dictionary. It also included the usefulness of a semi-bilingual EEUD. However, this paper presents analysis of one of the sections of the questionnaire in which the learners were asked to opine about the importance and usefulness of different types of dictionaries for their learning of the English language. The

unstructured interviews conducted immediately after administering the questionnaires to crosscheck the responses of the participants have also been made part of the analyses.

3.3.2 Questionnaire for Teachers of English

Another questionnaire was administered among the teachers to note their opinion on the use and usefulness of different types of dictionaries. The questionnaire designed for teachers for the main project consisted of four sections. However, the present paper reveals the results of the responses to only the section concerned with the teachers' preference for a specific kind of dictionary. The questionnaire also contained open-ended questions where the teachers were asked to give reasons for their recommendation of a particular type of dictionary. The analyses of the teachers' responses to the closed-ended questions have been shown in percentages that are supported by the qualitative discussion.

4. Results

The results discussed in the sections to follow are based on the analyses of data collected from participant students and teachers of the English language. The analyses of the questionnaire for students and teachers have been presented and discussed separately and conclusions are drawn by combining the results.

4.1 Learners' Questionnaire

The questionnaire administered for collecting data from the participants consisted of different questions related to the use and usefulness of the three types of dictionaries, namely, monolingual, bilingual and semi-bilingual. The data collected from the participants have been analysed and the results are presented below.

4.1.1 Learners' Need for More than One Dictionary

The participants were asked whether they used more than one dictionary in order to find the same kind of information. In response to this question, 76% (i.e. 303 out of 393) of the participants responded in positive. High percentage indicates that the learners of the English language in Pakistan require more than one dictionary to satiate their vocabulary needs.

4.1.2 Types of Dictionary(ies) used

The analysis reveals that participant students used all the three types of dictionaries in question. There were 340 students who used an EED (i.e. mono-lingual L2-L2 dictionary). There were 370 learners who said that they used an EEUD (i.e. semi-bilingual L2-L2-L1 dictionary). Then there were 216 students who told that they utilised an EUD (i.e. bilingual L2-L1 dictionary). The line graph (Figure 1) shows an analysis of the comparison of the number of participant learners using three types of decoding dictionaries.

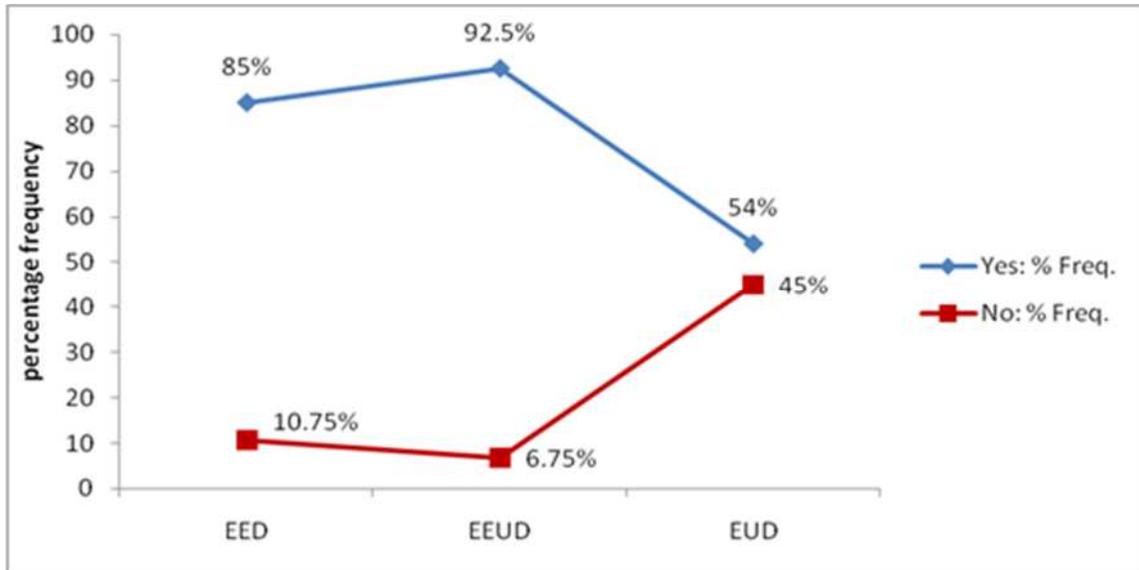


Figure 1. Comparison of the Number of Users of Decoding Dictionaries

The graph (Figure 1) illustrates a comparison of the number of participant users of decoding dictionaries. The frequencies shown in the graph depict the percentage taken out of the total number of participants, i.e. 400. The movement of graph line shows that majority of the participants used an EEUD to cater for their learning needs. Number of such users was more than 90% of the participants of this research. The monolingual EED was the next most favoured dictionary. A large number of participants, i.e. 85%, utilised this type of dictionary. A bilingual EUD was at the third place with 54% participant users of this type of dictionary. According to the participant learners, they have been using these types of dictionaries for a long period of time.

4.1.3 Frequency of Using Different Kinds(s) of Dictionary(ies)

One of the questions was concerned with seeking information about the frequency of using three types of dictionaries. In the form of a Likert¹ scale, the participant students were given five options, from 'daily' to 'not at all'. The response data about using three types of decoding dictionaries reveal that monolingual EED was daily used by 24% participants, and it was used twice a week by 28% respondents. As far as the semi-bilingual EEUD was concerned, a large number, i.e. 56% students made use of it daily, and 19% participant learners consulted it twice a week. On the other hand, a bilingual EUD was daily used by 15% of the respondents, while 6% used it twice a week, and 45% of the participants never used this type of dictionary. In order to draw a comparison of the frequency of the use of three kinds of dictionaries, the percentages of the responses are presented in line graph (Figure 2).

¹ The Likert scale, introduced in the 1930s and associated with Renesis Likert, is a rating scale used to measure the strength of agreement to the clearly worded statements, and to find answers to the questions in research.

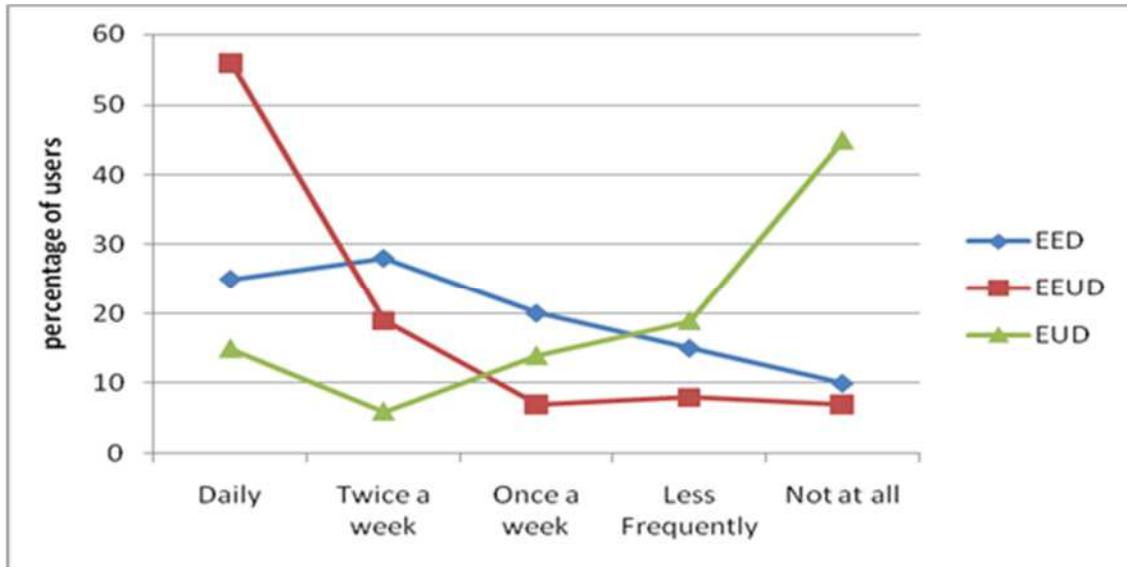


Figure 2. Frequency of the Use of Different Types of Decoding Dictionaries

The graph (Figure 2) illustrates the analysis of the frequency of the uses of different kinds of dictionaries. The lines show that the frequency of the use of semi-bilingual EEUD was on the highest level, followed by monolingual EED. On the other hand, the tendency shown in the graph reveals that a bilingual EUD was used far less frequently in comparison to the two kinds of other dictionaries. The figures show that the number of students who did not use a semi-bilingual or a monolingual dictionary at all was far less than those of the users of an EUD.

4.1.4 Extent of Usefulness of a Particular Type of Dictionary in Learners' Opinion

In order to find the usefulness of a particular kind of dictionary, the participants were asked to respond to the five options ranging from 'very useful' to 'useless'. The analysis of the participant learners' responses reveals that a majority of them believed that a semi-bilingual EEUD was useful for their studies. There were 90% (very useful=68%; useful=22%) participants who favoured this type of dictionary. Only 1% respondents considered it to be less useful for their studies. Next to the EEUD was a monolingual EED. There were 81% (very useful=50%; useful=31%) participants who considered it to be useful for their learning of English. Only 3% students considered it to be less useful, and none of the participants considered it to be useless for their learning.

On the other hand, the participants' views regarding a bilingual EUD showed that there were 40% (very useful=11%; useful=29%) users who believed it to be useful for their learning of the English language. 16% students believed that this type of dictionary was less useful. It is important to note that 43% participants did not respond about the usefulness of a bilingual dictionary. The analysis of comparison of the usefulness of three types of dictionaries for the participants is shown in Figure 3.

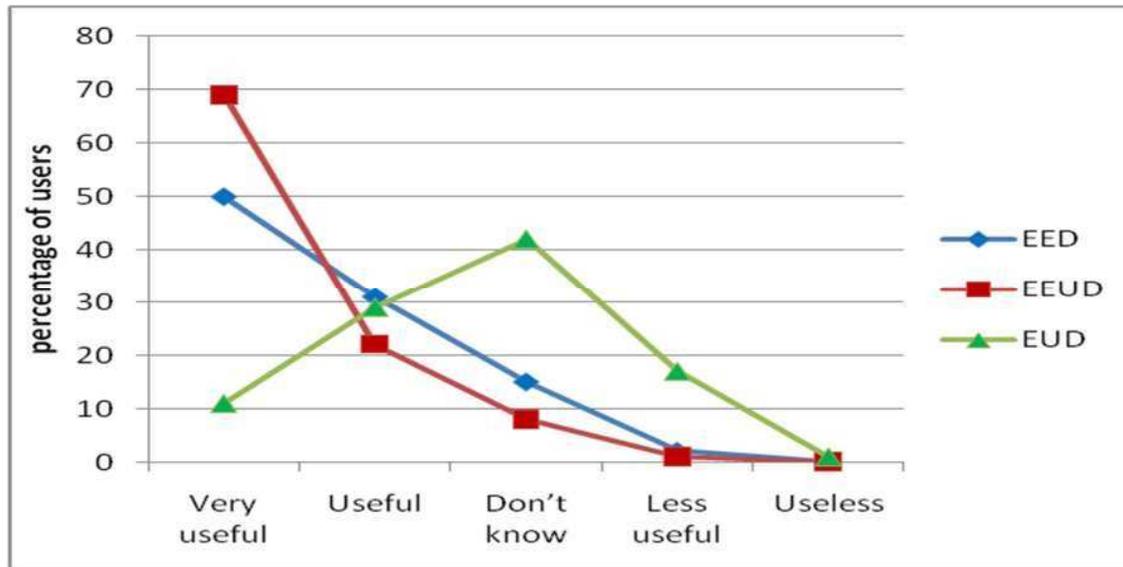


Figure 3. A comparison of students' views about the usefulness of dictionaries

The line graph (Figure 3) illustrates the participants' ratings based on their experience of using three types of decoding dictionaries. The graph reveals that 69% participants considered an EEUD to be very useful, while 50% learners believed that an EED was very useful for their learning of the English language. As far as a bilingual EUD was concerned, 11% participants considered it to be very useful. Considering the comparative analysis of the usefulness of three types of decoding dictionaries, an EEUD comes out to be the most useful dictionary.

4.2 Teachers' Questionnaire

This section of the paper presents analysis of the questionnaire data collected from twenty-five participant teachers. It shows teachers' choice and advice to their students to use particular type(s) of dictionaries. Teachers were also asked to give reasons of their choice. The analysis of the teachers' opinion is presented in two sections. The first section provides the analysis of their recommendation of a particular kind of dictionary. The quantitative results of their answers to the closed-ended questions are shown in the form of a line graph. The second part presents the qualitative descriptive analysis of the open-ended questions where the reasons of the teachers' choice of a particular type of dictionary have been discussed.

4.2.1 Teachers' choice/Recommendation of a Dictionary for their Students

Regarding the teachers' advice to their students to use a particular kind of dictionary, ten participant teachers recommended specifically an EED, and 11 teachers preferred only an EEUD, while 4 teachers recommended both EED and EEUD for their students. None of the teachers recommended an EUD. The responses of the teachers are shown in Figure 4.

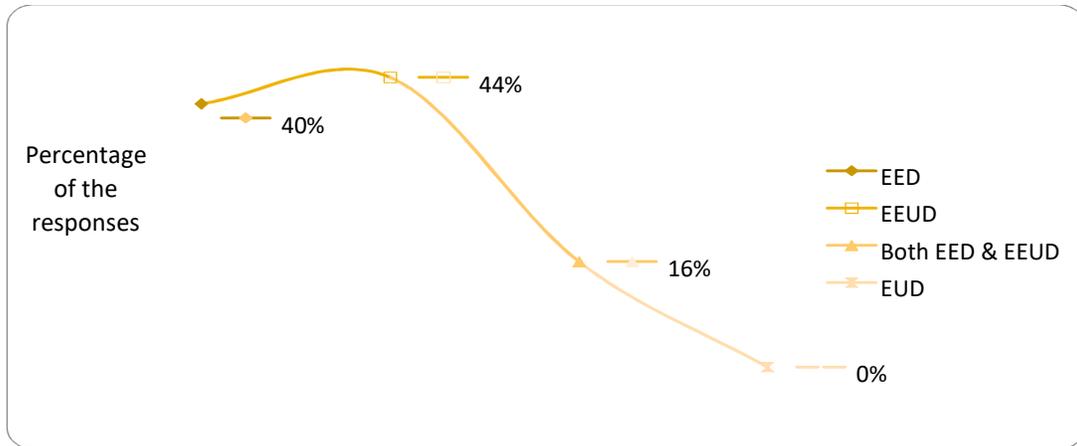


Figure 4. Teachers' preference of different types of dictionaries

The line graph (Figure 4) illustrates the choice of teachers for a particular type of decoding dictionary for their students. The graph shows that teachers mostly recommended two types of dictionaries for their students. The dictionaries preferred by them were semibilingual EEUD and monolingual EED. The percentage analysis presented in the graph shows that 44% responses were in favour of an EEUD, while 40% teachers chose EED. There were 16% of the participant teachers who recommended both EED and EEUD for their students.

4.2.2 Teachers' Reasons to Recommend a Particular Type of Dictionary

The teachers were requested to give reasons in support of their views why they recommended one or more type(s) of dictionary for their students. Their responses to the openended questions in favour of EED and EEUD are discussed in the following sub-sections.

4.2.2.1. Teachers' views about a monolingual EED

The participant teachers who advised their students to use a monolingual EED provided a number of reasons in support of their choice. Their views are summarised in the following lines:

A monolingual EED enhances learners' capability to learn the English language. It helps develop learners' habit to think in English. It gives information about the words with examples. It uses the standardised Received Pronunciation (RP) system, which improves the learners' skill in pronunciation. It is reliable in terms of meanings. It gives synonyms of words in English which improves students' vocabulary. Teachers also believed that the use of a monolingual dictionary paved the way to learning the English language without a reference to their L1.

4.2.2.2. Teachers' views about a semi-bilingual EEUD

The teachers who recommended students to use a semi-bilingual EEUD also provided reasons to support their point of view. The reasons given for their choice are summarised below:

In Pakistan, the Grammar Translation Method is predominantly exploited for teaching of the English language, and in this scenario a semi-bilingual EEUD can be more useful than a monolingual EED. A majority of the students in Pakistan call for the meanings of English words in both English and Urdu. As a semi-bilingual EEUD provides the meanings in both languages, it helps the learners to improve their understanding of English vocabulary. The learners who find it difficult to comprehend the meanings in English can understand words in Urdu. Thus, a semibilingual EEUD facilitates the learning of the English language.

5. Discussion

The analyses of the data show that learners of the English language in Pakistan used more than one dictionary. The two most preferred dictionaries were semi-bilingual EEUD and monolingual EED. A large majority of the 400 participant students made use of these types of dictionaries. Data also show that some of the participants had never used one or the other type of dictionary at all.

According to the participants' opinion, an EEUD was the most frequently used dictionary of all the three types of decoding dictionaries. After EEUD, the second most frequently used dictionary was a monolingual EED. A bilingual EUD was the least frequently used dictionary among the three dictionaries compared by the participant students of this study.

The comparison about the usefulness of the three types of decoding dictionaries shows that an EEUD was considered to be the most useful dictionary by the learners. Next to an EEUD was an EED. Majority of the users were contented with the different features of an EEUD like the structure and the provision of contents, and they also found it easy to use. A protocol study of the use of semi-bilingual EEUD conducted by Aslam, Iqbal and Choudhary (2016) also proved the usefulness of this type of dictionary.

Data collected from the teachers of the English language show that EEUD and EED are almost equally recommended. Most of the participant students used an EEUD and an EED as advised by their teachers, however, many of them also exploited an EUD to understand English words. Majority of the students regarded both EEUD and EED to be highly helpful for their learning of the English language.

6. Conclusion

Findings of the research reveal that, although learners of the English language exploit all three types of decoding dictionaries, a large number of them consider a semi-bilingual EEUD to be the best source of information for their studies. Along with an EEUD, the closest most preferred dictionary is an EED. The teachers' difference of opinion regarding their recommendation for a monolingual and semi-bilingual dictionary shows the difference of two schools of thought in Pakistan regarding the teaching of the English language. One of them supports the use of the Direct Method (DM) for teaching of the English language and believe that any attempt to teach English using other methods hampers the learning process. The supporters of Grammar Translation Method (GTM) think that the significance of reference to the learners' first language for teaching second or foreign language is of vital importance, and

it pertains to teaching of the English language as well. The responses of the participant teachers show the same difference of opinion. Followers of the GTM encourage and recommend the use of a dictionary which provides meanings in both the languages, i.e. English and Urdu. On the other hand, those who favour Direct Method for teaching of the English language recommend a monolingual dictionary as the basic source of information, and some of them also recommend an EEUD as an additional source for reference purposes. Majority of the learners follow their teachers' advice to use a dictionary for learning of the English language.

References

- Aslam, R. F. M., Iqbal, Z., & Choudhary, N.A. (2016). Protocol study of the use of semibilingual (English-English-Urdu) dictionary. *ELF Annual Research Journal*, 18, 103122
- Atkins, B.T.S., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. New York: Oxford University Press
- Barnhart, C. L. (1962). Problems in editing commercial monolingual dictionaries. In W. Householder, W. Fred, & S. Saporta (Eds.), *Problems in lexicography* (pp. 161-181). Bloomington: Indiana University.
- Bejoint, H. (1981). The foreign students' use of monolingual English dictionaries: A case study of language needs and reference skills. *Applied Linguistics*, 2(3), 207-222.
- Diab, T. (1990). *Pedagogical lexicography: A case study of Arab nurses as dictionary users*. Tübingen: Max Niemeyer Verlag.
- Hatherall, G. (1984). Studying dictionary use: Some findings and proposals. In R. R. K. Hartmann (Ed.), *LEXeter '83 Proceedings: Papers from International Conference on Lexicography at Exeter, 9- 12 Sept. 1983, Lexicographica Series Maior 1*, (pp. 183-189). Tübingen: Niemeyer.
- Hartmann, R. R. K. (1983). The bilingual learners' dictionary and its uses. *Multilingua*, 2(4), 195-201.
- Hartmann, R. R. K., & James, G. (comps.). (2002). *Dictionary of Lexicography*. Taylor & Francis e-Library
- Iqbal, Z. (1987). *Aspects of the learner's dictionary with special reference to advanced Pakistani learners of English* (Unpublished doctoral dissertation). University of Aston, Birmingham.
- Kernerman Dictionaries. (N.D) *Kdictionaries*. Retrieved October 24, 2009, from <http://www.kdictionaries.com/>
- Kipfer, B. (1985). *Dictionaries and the intermediate student: Communicative needs and the development of user reference skills* (Unpublished M. Phil. Thesis). University of Exeter, Devon UK.
- Laufer, B., & Hadar, L. (1997). Assessing the effectiveness of monolingual, bilingual, and "bilingualised" dictionaries in the comprehension and production of new words. *The Modern Language Journal*, 81(2), 189-196.
- Laufer, B., & Melamed, L. (1994). Monolingual, bilingual and 'bilingualised' dictionaries: which are

- more effective, for what and for whom? In W. Martin, W. Meijs, M. Moerland, E. ten Pas, P. van Sterkenburg, & P. Vossen (Eds.), *EURALEX 1994 Proceedings* (pp. 565-576). Amsterdam: Vrije Universiteit
- Lew, R. (2002). Questionnaire in Dictionary Research: Reexamination. In A. Braasch, Anna & C. Povlsen (Eds.), *Proceedings of the Tenth EURALEX International Congress, EURALEX 2002, 1* (pp. 267-271). Copenhagen: Center for Sprogteknologi (CST). Retrieved April 01, 2009 from www.staff.amu.edu.pl/~rlew/pub/Lew_2002b.doc
- Lew, R. (2004). *Which dictionary for whom? Receptive use of bilingual, monolingual and semi-bilingual dictionaries by Polish learners of English*. Poznan: Motivex
- Marello, C. (1998). Hornby's bilingualised dictionaries. *International Journal of Lexicography*, 11(4), 292-314.
- Nakamoto, K. (1995). Monolingual or bilingual, that is not the question: The 'bilingualised' dictionary. *Kernerman Dictionary News 2*. Retrieved October 24, 2009, from <http://kictionaries.com/newsletter/kdn2-2.html4>.
- Quirk, R. (1973). The social impact of the dictionaries in UK. In R. I. McDavid, & A. R. Duckert (Eds.), *Lexicography in English* (pp. 76-88). New York: NY Academy of Sciences.
- Singh, Y. K. (2005). *Research methodology*. New Delhi: A.P.H. Publishing Corporation
- Sterkenburg, P. V. (2003). *A practical guide to lexicography*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- Thumb, J. (2004). *Dictionary look-up strategies and the bilingualised learner's dictionary: A think-aloud study*. Tübingen: Max Niemeyer Verlag
- Tomaszczyk, J. (1979). Dictionaries: Users and uses. *Glottodidactica*, 12, 103-119.
- Zgusta, L. (1971). *A manual of lexicography*. Prague: Academia Publishing House of Czechoslovak.

Process of Terminography of Information and Communication Technology (ICT) for Students in the Urdu Language

Masroor Sibtain¹, Hafiz Muhammad Qasim², Nazia Suleman³

masroorsibtain@gmail.com, m_qasim@hotmail.com, naziasuleman@ciitvehari.edu.pk

¹Assistant Professor, Department of English, The Women University Multan

²Lecturer, Department of Applied Linguistics, Government College University Faisalabad

³Assistant Professor, Department of Humanities, COMSATS Institute of Information Technology, Vehari

Abstract

The purpose of this study is to investigate the word-coining process for the Urdu language particularly within the genre of Information and Communication technologies (ICT) in Pakistan for proposing the secondary terms. Overreliance on borrowing and coin formation used extensively for term formation in Urdu is not only cognitively challenging but has also painstakingly generated Urdu terms with widely accepted pronunciation which also triggers schemata to the ICT terms. Coinages and loan translations from the root languages are in abundance in case of the Urdu language. Cabre's (2000) Communicative Terminology Theory (CTT) has been used as a theoretical framework to coin new Urdu terminologies. In Pakistan, the need to coin new terms is emerging most strongly after the verdict of Supreme Court of Pakistan in 2014 to promulgate the use of Urdu (Target Language) as an official language at the national forum. Keeping in view the situation, methods from the historical study of term coining have been opted to explain the coining process. This study concludes that the redundant coining and loans from English tarnishes the purity of the Urdu language. This is not only limited to ICT terminology but can be further extended in all domains with a lexical debt owed to the English language in its day to day linguistic parlance. Borrowing, loan translation, hybrid formation, and proper translation are some popular means employed as methods of term translation.

Keywords: coining, Muqtadira, national language, target language (TL)

1. Background of the Study

Etymologically the vocabulary of any language is compiled from two assemblages: lexical items from the native language and the loanwords which are foreign elements. Haspelmath et al. (2009) have succinctly worded the omnipresence of the loanwords in the languages around the globe and stated that languages 'entirely devoid of loanwords' is a rare phenomenon. At the same place, the transference of the concepts from one place to another is a substantial process as Postolea (2012) opines that 'concepts, being the basic unit of knowledge, are almost never bound to their place of birth. Packed in their original language, they easily travel across borders'. In the ICT word stock of the Urdu language, an overriding ascendancy of borrowings from English is witnessed in Urdu, a language struggling hard to preserve its purity against the linguistic pressures inherent to globalization and technological developments.

Sager (1990) has segregated the definitions for the primary as well as secondary term formation process. The first process may be named after the name of its creator or another in the native language while in secondary term formation it involves certain processes of coining or borrowing. Within the genre of ICT, almost all the imported terms originated in the English language, thus making English the original language of creation. In the Pakistani context, term formation process is a complex one, and it involves a gamut of strategies which include coinage, borrowing, derivation, and affixation, etc.

2. Introduction

The modern age has witnessed an exponential growth in all branches of knowledge, human sciences, and social sciences, and new concepts are emerging in every field of knowledge which requires the formulation of terms for communication. In the postmodern era, diversity in all kinds of knowledge has got an unprecedented upsurge. The era in which we are living is of information and knowledge which poses challenges in many fields of communication, and terminology is one such field. The challenges we encounter today include an exponential growth in the number of specialized documents in which terms and new concepts are introduced. Wüster in his 1976 work (cited in Campo, 2012) considers terminology as being located at the intersection of linguistics, logic, ontology, information science, computer science, and individual disciplines.

The West, however, is continuously expanding its terminology bank with the expansion of branches in sciences and technology. The West as such has devised the ways to meet the challenges ahead and to pattern the life to seek the solutions pertaining to terminology. It is a rough estimation that approximately twenty-five new terms are daily added to the one million existing terms. The question arises as to how this large number of new terms will be conveyed in books and dictionaries in approximately two hundred and fifty disciplines in the mammoth domains of knowledge. However, the fact remains that the West is also meeting the task with not much ease as they have to borrow words from Persian, Sanskrit, Chinese, Japanese, etc. Hardly any language can claim to be self-sufficient in reference to terminology. Durrani (1993) elaborates the following facts to support the stance that English is not the only language to use. It is an estimation that each year 2 million publications appear and 50 to 60 thousand periodicals are brought out all over the world (cited in IOUTN, 1985). Out of it, 51% are published in English, and it is impossible to get translation of other 49% in English. The International Translation Centre in Delft has been able to translate only 4% of the remaining material into English. So, English is providing us with only 55% of the existing knowledge.

Language is categorised as integrative as well as divisive. The decision-makers use it either way (Rahman, 2005). The Urdu language has been lucky in one respect: it began to attract the attention of terminologists while in its infancy. The various arts and sciences do have rich Urdu terminologies which are because of its borrowing capacity as well as its comprehensiveness. Durrani (1993) elucidates that a large number of Persian-Arabic terminologies have been absorbed in the Urdu language.

The study has discussed in detail the term formation methods employed by the researcher to propose new terminologies in ICT. Methodology prescribed and practiced by Muqtadira Qaumi Zabaan (1986) has been employed to meet the scientific needs of the time. Muqtadira translated Windows in the Urdu language in 2010 with the joint collaboration of Microsoft (<http://cle.org.pk/>). The root languages particularly Persian, Arabic, and Turkish have been consulted to coin new terms. Various academies, societies, departments, and commission around the globe such as Academy of Persian Language and Literature (APLL), Hellenic Society for Terminology (ELETO), National Language Promotion Department

(NLPD), and Commission of Scientific & Technical Terminology (CSTT) have been established which are working to preserve their national languages.

The study mainly investigates the following areas from the perspective of the Urdu language.

1. Is there any rule in the history for translating terms from foreign language into Urdu?
2. If yes, then can it be applied on the ICT terms to coin Urdu equivalents?

3. Research Methodology

This is a descriptive study in Urdu terminology. It analyzed the terms that were created using existing rules in the Urdu language. This study focused on the strategies employed by the subject experts in the domains of Information and Computer Technology and Diplomacy to propose new equivalents from the existing sources of the Urdu language as well as from the roots of the Urdu language. The methodology adopted for the current study is the one prescribed by Urdu Linguist Moulvi Abdul Haque and Muqtadra Quomi Zuban in 1986 in order to coin and finalize the terminologies.

The study followed a list of English ICT terms, along with their equivalents in the target language (TL) and their phonetic representations. Glosses are elaborated in the form of a table to get the comprehension of the terminology. The justification from the history has also been given to justify the proposed terminology. While scrutinizing the list, the ICT experts were also enquired to seek their viewpoint about the proposed terminologies. This paper cites initially fourteen frequently used ICT terms against which the Urdu terms have been proposed. The terms comprised of the name of branch, names of applications, and significant terms of the subject. The proposed terminologies have either been borrowed from the root languages or coined based on the existing sources of the Urdu language.

4. Theoretical Framework of the Study

Various theories relevant to terminography have been proposed by the terminologists from time to time. Theorists, while describing alternatives, favour that generalization would be applied to all other language situations (Hadebe, 2002). To support the aforesaid view, theories including L'Homme's (2003), the Communicative Theory of Terminology (Cabr  Castellv , 2003), Sociocognitive Terminology (Temmerman, 2003), and Cognitive-based Theory of Terminology (Kageura & Umino, 1996) are often cited. The above-mentioned theories arose in the wake of Wuster General Theory of Terminology (GTT), which is prescriptive in nature and has an inclination towards communicative, social, and cognitive aspects of terminology. These theories are descriptive as they take into account different terminological activities carried out under different conditions. This paper also encompasses, in general, the linguistic principles of term formation, exploring the methods to coin new terms in the Urdu language. So, the Communicative Theory of Terminology (CTT) by Theresa Maria Cabre was followed as a framework to provide guidelines for the term formation in the Urdu language. The principles contained in the theory enabled the researcher to formulate the term formation strategies for the Urdu language in the field of ICT.

Using a language activates its linguistic resources as the study in hand. Cabré (2003), who suggested Communicative Theory of Terminology (CTT), which is descriptive in nature, opines that theory can never be prescriptive because a theory is the combination of a unit where axioms are integrated coherently. Cabré, Condamines, and Ibekwe-San Juan (2005) is of the view that “this theory is characterized by a strong orientation towards practical problem solving and establishing methods with scientific justification in order to make problem solving more efficient.” To this extent, CTT has paved the way to a research on multiple aspects of terminology as conceptual and terminological relations, term creation and the application of varied linguistic models to the study of terminology. So, it can be derived that CTT is developed from a more holistic and integrative point of view. Its main focus lies in the transfer of knowledge which is the communicative aspect of terminology. Moreover, CTT mainly focuses on the transfer of knowledge, which is the communicative side of terminology. All the professions dealing with special knowledge require terminology which ultimately leads to documentation, language planning, and language engineering. The next section elaborates the Muqtadira’s rules for terminology coining in the Urdu language.

5. Muqtadira’s Rule of Terminology Establishing and Coining

The main reason of the establishment of Muqtadira Qaumi Zuban in 1979 was to promulgate the use of Urdu as the official language in academics and to propose language policies to the Government of Pakistan. Terminology did not fall directly within its ambit, but as the educational material encompasses certain terms, so they have to work upon it as well.

Different trends remained prevalent at different time and span to coin the terminologies, and a series of seminars were organized by the Muqtadra Qaumi Zuban to lay a foundation for devising rules for terminology.

Jalibi (1991) devised some rules of terminology coining. He laid a foundation of English-Urdu dictionary and the maximum part of it comprises of terminographic rules. Clauses of terminology for coining adaptable terms as discussed by Jalibi are as follows:

- 1) International terminologies should be transliterated like Magnesium (Mg), nicotine and malaria, etc.
- 2) Chemical formulas and symbols will remain in English.
- 3) Technical names of botany and zoology should be placed with Urdu equivalents or it should have terminology otherwise it should be transliterated in Urdu. Like سا بل تا زشه (Sabalpalmete).
- 4) If a word is not available in Urdu language but can be adjusted according to Urdu language, then it should be taken in Urdu like: Cabala كبا لا, Cabalian كبا لیت, cabalistic كبا لیت etc.

After compiling a list of ICT terms which needs equivalents, the equivalents were proposed from the Urdu sources with the consent of ICT experts.

6. Discussion about Proposed Terminologies

The lack of terminology can be compensated if all the existing sources are mined in the best possible manner. Few terms have been selected randomly from the field of ICT to coin Urdu equivalents. Terminologies of ICT are shown in the Tables below.

Table 1

Terminology for Analogue Computer

English Term	Analogue Computer		
POS Tagging	Noun		
Description/Definition	A machine or electronic circuit designed to work on numerical data represented by some physical quantity (e.g. rotation or displacement) or electrical quantity (e.g. voltage or charge) which varies continuously, in contrast to <u>digital</u> signals which are either 0 or 1.		
Proposed Equivalent Urdu Term	تما ٹلی کمپیوٹر	Etymology aspect of Proposed Term	English+Arabic
Phonetics	ʈɑɲʂɑlʲi kɑmpö tə	Gloss	Tamasli computer
Formation Rules	As described by Delhi College Terminology Principles, the famous words need to be picked as it is in the Urdu language. So, computer has been transliterated while تما ٹلی is borrowed from the Arabic language.		

The coined term is the combination of two terms which are also from two languages. As computer is the most common word amongst the Urdu speaking community, it has been transliterated while analogue has been given an equivalent from the Arabic language.

Table 2 elaborates the Urdu elaboration of *Application* as it is the most frequently used term in ICT.

Table 2

Terminology for Application

English Term	Application		
POS Tagging	Noun		
Description/Definition	the special use or purpose to which something is put: a technology having numerous applications never thought of by its inventors.		
Proposed Equivalent Urdu Term	برنامہ	Etymology aspect of Proposed Term	Persian
Phonetics	barnamə	Gloss	Barnama
Formation Rules	This word has been borrowed from the Persian language. The inclination towards Persian is more dominant in every phase of Urdu terminology coining.		

The proposed terminology is borrowed from the Persian language but is quite familiar for the Urdu speakers as all its parts are the ingredients of the Urdu language as well.

Table 3 depicts the term *Automation* with its gloss and defined prospects.

Table 3

Terminology for Automation

English Term	Automation
POS Tagging	Noun
Description/Definition	the technique, method, or system of operating or controlling a process by highly <u>automatic</u> means, as by electronic devices, reducing human intervention to a minimum.
Proposed Equivalent Urdu Term	خودکار سازی Etymology of the proposed term Urdu Gloss Khudkarsazi
Phonetics	khodka:rsa:31
Formation Rules	Here the equivalent is sort from the common use Urdu. Haq as discussed by durrani (1993) that the priority should be given to those terms which are easy to pick from the Urdu language.

The rules defined by Haque have been applied to propose the term from the Urdu language. The existing sources of the Urdu language helped to propose this term. For *auto* the word has been chosen from the existing sources and has been used in ICT as well to fulfill the need of communication.

Table 4 explains the formation rules as well as the POS tagging of *Autodetect*.

Table 4

Terminology of Autodetect

English Term	Autodetect
POS Tagging	Noun
Description/Definition	self-diagnosing functionality of a software device/ application.
Proposed Equivalent Urdu Term	خودکار تشخیص Etymology Urdu Gloss Khudkartakhshees
Phonetics	khodka:rtakhfi:s
Formation Rules	Sorted from the existing vocabulary of the Urdu language.

The term is split into two parts. The translation once proposed for auto has been used here to propose the term. The existing vocabulary has been used and is in alignment with the definition.

Table 5 proposes the secondary term formation of *Autodetection* which is a noun. Here the literal meanings of the term have been translated.

Table 5

Proposed Terminology for Autodetection

English Term	Autodetection
POS Tagging	Noun
Description/Definition	The property of self-diagnosing functionality of software device application
Proposed Equivalent Urdu Term	خودکار تشخیص سازی Etymology of the Proposed Urdu Term
	Gloss Khudkartakhsheessazi
Phonetics	khodka:rajka:rsazi:
Formation Rules	As per the principles of Delhi College, it has been coined from the Urdu sources which fits well with English terminology.

The proposed terminology justifies the definition and also it is a suitable equivalent as here the extension of one term is forwarded.

Table 6 provides the proposed term for *Secondary Storage* in Urdu.

Table 6

Proposed Terminology for Auxiliary Storage

English Term	Auxiliary Storage
POS Tagging	Noun
Description/Definition	<u>Secondary storage.</u>
Proposed Equivalent Urdu Term	اضافیتخزین Etymology of the Proposed term Persian
	Gloss Izaafitakhzin
Phonetics	iza:fitakhzi:n
Formation Rules	In terminology coinage here, both the words are chosen from the Urdu language's existing sources which explain the definition well.

The terminology is picked from the root language Persian and satisfies the definition and is equally acceptable for the computer scientists.

Table 7 elaborates the term for *Buffer* with all the possible dimensions needed to a terminologist in the Urdu language.

Table 7

Terminology Proposed for Buffer

English Term	Buffer
POS Tagging	Noun
Description/Definition	A device or an area of a computer that temporarily stores data that is being transferred between two machines such as a computer and a printer.
Proposed Equivalent Urdu Term	ٹامفون Etymology aspect of the Proposed Term Turkish Gloss Tampfoon
Phonetics	ta:mfon
Formation Rules	The word is borrowed from the Turkish language. In Turkish, it is written as <i>tampon</i> . As per term formation rules the p of Turkish is replaced by F sound in Urdu.

Buffer is commonly used term in ICT so it has been borrowed from the Turkish language.

Table 8 highlights the terminological aspects of *Cache*. The formation rules for cache are the combinations of the two methods although both the terms are the part of Urdu language.

Table 8

Terminology Proposed for Cache

English Term	Cache
POS Tagging	Noun
Description/Definition	a temporary storage space or memory that allows fast access to data: <i>Web browser cache</i>
Proposed Equivalent Urdu Term	نہانگاہ Etymology aspect of the Proposed Term Urdu Gloss Nehanga
Phonetics	neha:ngah
Formation Rules	Here the principle of Revision has been used as suggested by Haque(19). He is of the view that new terminologies can be coined by the use of existing terminologies so from میان گاہ it has been coined as نہانگاہ.

Table 9 offers the terminology for *Chip*, one of the most frequently used term in ICT as well as electronics.

Table 9

Terminology Proposed for Chip

English Term	Chip
POS Tagging	Noun
Description/Definition	a circuit of transistors, resistors, and capacitors constructed on a single semi-conductor wafer or chip, in which the components are interconnected to perform a given function.
Proposed Equivalent Urdu Term	تراشه Etymology of the Proposed term Urdu Gloss Tarasha
Phonetics	ta:raʃə
Formation Rules	Delhi College Principle has been applied here. The equivalent has been given from the daily use words.

Table 9 proposes terminology for the branch of Computer Science, which is a combination of an Urdu and Persian equivalent.

Table 10

Terminology Proposed for Cryptography

English Term	Cryptography
POS Tagging	Noun
Description/Definition	the science or study of the techniques of secret writing, especially co deand cipher systems, methods, and the like.
Proposed Equivalent Urdu Term	رمزنویسی Etymology of the Proposed term Persian+Urdu Gloss Ramznvisi
Phonetics	ra:mznavi:si
Formation Rules	Dr. Abdul Rehman has predefined specific terms for few often used words in English like <i>Graph</i> and <i>phone</i> . The term رمز is borrowed from the Persian language.

Table 10 proposes the term for the word most frequently used in ICT as well as in many other domains of knowledge.

8. Conclusion

This study is a significant contribution towards enriching the Urdu language. Perhaps, the first thing to consider is language engineering not language planning. Due to the foreign origin of the approximately all scientific concepts, the academicians are left only with two choices: either using the foreign terms or selecting new equivalents in the Urdu language to make it comprehensible to the users. There is no doubt that using the foreign terms is far easier and convenient, but consequently the Urdu language will be full of foreign terms only and nothing will remain out of it. Moreover, these foreign terms cannot be used by the Urdu speakers to make new combinations and constructions. Keeping in view the facts, we can say

that there is a dire need to carry research on the scientific terms in order to enrich the treasury of the Urdu language with native and local equivalents to express new concepts.

References

- Cabré, M. T., Condamines, A., & Ibekwe-SanJuan, F. (2005). Introduction: Application driven terminology engineering. *Terminology. International journal of theoretical and applied issues in specialized communication*, 11(1), 1-19.
- Cabré, C. M. T. (2003). Theories of terminology: Their description, prescription and explanation. *Terminology*, 9(2), 163–199.
- Cabré, M. T. C. (2000). Elements for a theory of terminology: Towards an alternative paradigm. *Terminology. International Journal of Theoretical and Applied Issues In Specialized Communication*, 6(1), 35-57.
- Campo, A. (2012). The reception of Eugen Wüster's work and the development of terminology, 378.
- Durrani, A. (1993). *Urdu Istalahat Nigari*. Islamabad: Muqtadra Qaumi Zaban.
- Hadebe, S. (2002). The standardisation of the Ndebele language through dictionary-making. *Harare: Alex Project, University of Zimbabwe*.
- Haspelmath, M. (2009). Lexical Borrowing: Concepts and Issues. *Loanwords in the world's languages: A comparative handbook*, 35-54.
- Jalibi, J. (1991). *Farhang Istalahat Jamia Usmania*. Vol. 1. Islamabad: Muqtadra Qaumi.
- Kageura, K., & Umino, B. (1996). Methods of automatic term recognition: A review. *Terminology*, 3(2), 259–289.
- L'Homme, M. C., & Bae, H. S. (2003). A Methodology for Developing Multilingual Resources for Terminology, 1–6.
- Postolea, S. (2012). From regularities to norms in the secondary formation of Romanian ICT terms. *Linguaculture*, 2, 101-115.
- Rahman, T. (2005). Passports to privilege: The English-medium schools in Pakistan.
- Sager, J. C. (1990). *A Practical Course in Terminology Processing*. Amsterdam: John Benjamins Publishing Company.
- Temmerman, R. (2003). Innovative methods in specialised lexicography. *Terminology*, 9(1), 117–135.
- Wuster, E. (1976). *Grundbegriffe bei Werkzeugmaschinen (in German) (1. Auflage ed.)*. London: Technical Press/Economic Commission for Europe of the United Nations.

A Corpus-Based Study of Nationality-Related Swearwords in British National Corpus

Ilyas Babar Awan¹, Muhammad Hussain²

¹Lecturer in English (PhD Scholar), Translation & Interpretation Department, National University of Modern Languages, Islamabad

²PhD English (Scholar), Faculty of Language & Literature, International Islamic University, Islamabad

Abstract

Swearwords are a part of everyday language use. A number of corpus-based studies such as gender, age, and social class have been conducted; however, nationality-related swearwords are not explored particularly with reference to British National Corpus (BNC). The aim of the present study is to investigate the use of selected swearwords that provide evidence for ethnic/national biased approach of one nationality for the other. The present study identifies and classifies selected swearwords by using BNC in order to find out occurrences of the target words, frequencies and collocations. It further traces out the differences between their use in spoken and written expressions. The results of the study show that British native speakers use swearword 'mick' against Irish nationals frequently which is 15.22 per million words and the frequency of 'Guido' which is used for Italian in BNC is 4.81 per million words. Furthermore, in BNC spoken expression the second highest number of swearword 'Paki' is against Pakistani nationals which is 1.06 per million words. Future researchers may further explore the BNC in different perspectives.

Keywords: corpus, nationality, swearwords, British National Corpus (BNC)

1. Introduction

The present paper explores the use of nationality-related racial slurs in terms of swearwords with reference to British National Corpus (BNC). The paper aims to investigate how various nationalities have been stereotyped by using frequent racial slurs against them in BNC. The study may be helpful to understand the western biased approach against nationalities across the world. The study may serve as pattern study to explore BNC in various perspectives that have received little attention of researchers. Besides, the article may also be a contribution in the domain of the empire writes back. Swearwords are the category in linguistic items which represent and channelize the anger, frustration, disgust, and hatred (Fägersten, 2012; Dewaele, 2015). By and large, it is a common characteristic of the world languages that they contain swearwords. Swearing words reflect the attitude of language users toward the opposite gender, class, relation, and nationality (Chen, 2004; Ljung, 2010). Some commonly used sex-oriented swearing words are not too old in the English language, we find their traces in medieval ages, and they are used in multiple connotations (McEnery, 2004; Montagu, 1967). Languages borrow lexical items from other languages as a natural process of cultural invasion; therefore, they also incorporate swearing words in them. With the changing socio-political scenarios, the use of nationality-based swearing words has been a contemporary phenomenon that needs to be further investigated in language. This study is a quantitative and systematic account of swearwords in the English language as found in BNC. British National Corpus (BNC) is a 100,000,000-word corpus that provides empirical evidence of the use of swearwords against particular nationalities as well as cultural and ethnical groups. The selected swearwords have been identified in multiple ways as collocation and concordance. By comparing certain swearwords, we have associated socio-

political and cultural perspectives with them. The aim of this study is to analyse the use of swearwords in written and spoken genre.

2. Literature Review

This part of the study is an extension of the previous research works in the area of swearwords in general and BNC in particular. This study is about the use of nationality-based swearwords and how they reflect the resistance and hatred against particular nationalities, as revealed in the British National Corpus. Besides, bulk of studies (see Popușoi, Havârneanu, & Havârneanu, 2018; Zenner & Ruette, 2018; Harismayanti, Budasi, Lin & Ramendra, 2017; Arsana, Putra, & Ramendra, 2017; Khoirunnisa & Wijayanto, 2017; Stenström, 2017; Rullyanti, 2017; Marom, 2014; Kwon & Cho, 2016; Güvendir, 2015; Kurniawan, 2015; Fägersten, 2014; Christie, 2013; Praschinger, Pomikal, & Stieger, 2011; Gauthier, 2012; Fras, 2012; Kirk, 2013; Suyanto, 2010; Hjort, 2009; Dewaele, 2004, among others) have been conducted on swearwords in various perspectives in general; however, the BNC has not been explored in terms of swearwords so far.

2.1. What Constitutes a Swear Word?

Swearing words are also known as offensive or profane words. Language users' associate expletive attributes while decoding curse, hatred, anger, and bias (Goddard, 2015; Moore, 2102).

These swear words unfold the use of language in the following manners:

- a) Through expletive attributes a speaker may indicate his irritation, anger, and occasionally approval.
- b) Sometimes the use of foul words, such as *fucking* and *goddamn* are used to put emphasis on certain feeling or to showcase frustration.
- c) Swearing words also reflect the change in age and youth's communication to deconstruct social interaction (Thorne, 2014).

2.2. Types of Swearing Word

Though the swearing words are considered bad language but no language of the world is swearing free (Darma, Wennyta, & Fitri, 2017). These words are categorized as:

- a. Swear words related to religion (e.g. Jesus, heaven, hell and damn)
- b. Swear words related to sex (e.g. f**k and c**t)
- c. Swear words related to racism (e.g. nigger)
- d. Swear words related to defecation (e.g. sh*t and pi**)
- e. Swear words related to homophobia (e.g. queer)
- f. Swear words related to nationality (e.g. paki, mick)

2.3. Functions and Motifs of Swearwords

Swearwords can be categorized in a number of ways and in accordance with different variables. One way of categorising swearwords is based on their motifs: religious, scatological, genital, intercourse, sodomy and mother (Ljung, 2006, p. 62). Another classification can be sex, religion, bodily functions, ethnic groups, food, dirt and death (Andersson & Trudgill, 1990, p. 55). Yet another order of grouping can be cursing, profanity, blasphemy, taboo, obscenity, vulgarity, slang, epithets, scatology, insults, and slurs (Jay, 1992, p. 1-8). A swearword, or phrase, can belong to more than one group at one and the same time like *fucking cunt*, which belongs to both the intercourse group and the genital group (Ljung, 2006, p. 63).

This study explores the use of swearwords in terms of nationality. It aims at finding out the frequencies of these words (Leech & Rayson, 2014) in overall corpus at first phase of the analysis which is followed by the second phase of analysis in which swearwords used in spoken corpus will be identified and their frequencies will also been mentioned. At the last phase of the analysis, the written texts will be analyzed in terms of swearwords used in order to express hatred for other nationalities.

3. Methodology

The aim of the present study is to explore the use of swearwords against nationalities with reference to BNC. For this purpose, various swearwords have been identified from various online sources (<https://www.noswearing.com/dictionary>). Having identified various swearwords, British National Corpus was used in order to find out frequent swearwords in terms of written and spoken perspectives. Overall swearwords were identified at the first phase of the analysis, followed by the identification of swearwords in the spoken and written corpus. Additionally, hits, total texts, words and frequencies were explored, identified, and classified in spoken and written data. The study employs quantitative and qualitative methods in order to investigate the selected swearwords by finding out the frequencies followed by interpretation of the data. Furthermore, corpus linguistics is itself termed as a fast-growing methodology in contemporary linguistics (Gries, 2009), thus frequencies and analysis have been carried out with the help of BNC results.

4. Analysis/Findings

The BNC data has been analyzed in this section. Our study has mainly focused on the words showing disrespect or hatred against a particular nationality. From BNC corpus, we have obtained the following results about the selected swearwords.

Table 1

Swearwords in Overall BNC Data

Swearwords	Nationality	Hits	Text/total Text	Words	Frequency
Mick	Irish	1496	300/40,48	98,313,429	15.22 instance per million word
Guido	Italian	473	39	Do	4.81
Pollock	Polish	267	95	Do	2.72
Spook	White	35	26	Do	0.36
Paki	Pakistani	28	18	Do	0.28
Spick	Mexican American	24	23	Do	0.24
Coon	African American	17	14	Do	0.17
Wop	Italian	14	10	Do	0.14
Wetback	Mexican	1	1	Do	0.01
Ruski	Russian	1	1	Do	0.01
Heeb	Jewish	0	0	0	0
Porch monkey	African American	0	0	0	0

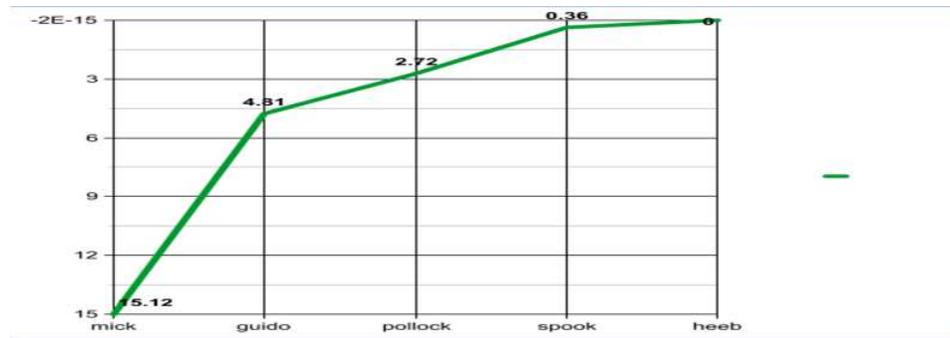


Figure 1. Quantitative findings of selected swearwords form Overall BNC Data
<https://corpus.byu.edu/bnc>

Table 2

Swearwords in Spoken BNC Data

Swearwords	Hits	Words/total Words	Words	Frequency
Mick	258	63/908	10,409,858	24.78 instance per million word
Paki	11	5	Do	1.06
Pollock	3	1	Do	0.29
Spick	2	2	Do	0.19
Spook	2	2	Do	0.19
Guido	1	1	Do	0.1
Porch monkey	0	0	0	0
Wetback	0	0	0	0
Heeb	0	0	0	0
Ruski	0	0	0	0
Coon	0	0	0	0
Porch monkey	0	0	0	0

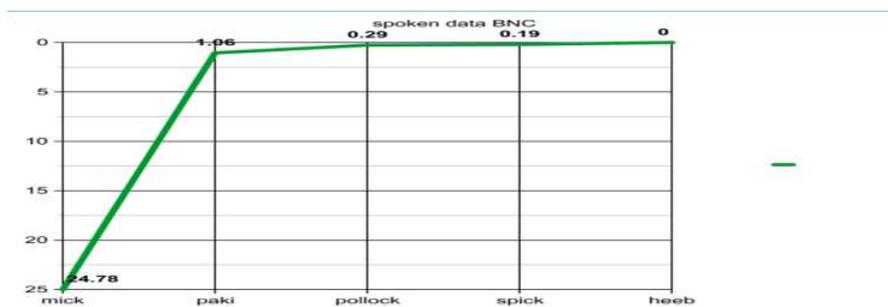


Figure 2. Quantitative findings of selected swearing words from spoken BNC data <https://corpus.byu.edu/bnc>

Table 3

Swearwords in Written BNC Data

Swearwords	Hits	Text/total text	Words	Frequency
Mick	1238	237/3140	87,903,571	14.08 instance per million word
Guido	472	38	Do	5.37
Pollock	267	95	Do	2.72
Spook	33	24	Do	0.38
Spick	22	21	Do	0.25
Paki	17	13	Do	0.19
Coon	17	14	Do	0.19
Wop	10	9	Do	0.11
Wetback	1	1	Do	0.01
Ruski	1	1	Do	0.01
Heeb	0	0	0	0
Porch monkey	0	0	0	0

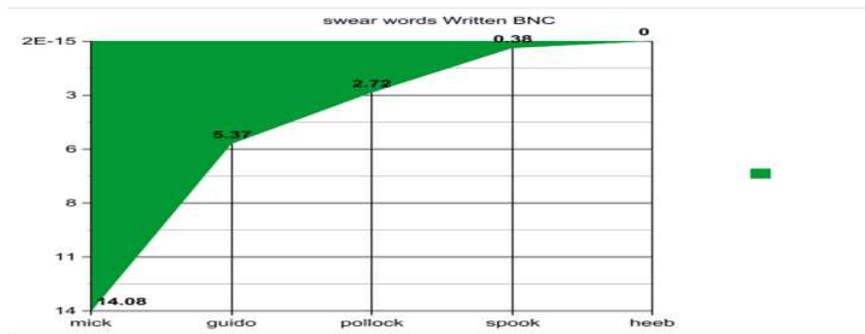


Figure 3. Quantitative findings of selected swearwords from written BNC data

<https://corpus.byu.edu/bnc>

5. Socio-Political Perspective Behind Swearwords

The purpose of collecting the data of nationality-based swearwords from BNC is to provide evidence against the British community that claims to be a multicultural community and pays respect to other cultural and ethnic groups. The result of Figure1 shows that British native speakers use most frequent swearword against Irish national (1496 hits from 300 texts). The frequency of the swearword *mick* (Irish) is 15.22 per million words. The second highest swearword used against Italian is *guido* (473 hits out of 39 texts). The frequency of *guido* in BNC is 4.81 per million words. The third highest swearword is used against Polish, which is *pollock* (267 hits out of 95 texts). The frequency of *pollock* in BNC is 2.72 per million words. Anglo-Irish relationships with respect to the swearword *mick* against Irish may be evaluated under the light of BNC. The Act of Union between these two cultural and national identities came into effect on January 1, 1801.

That, from the first day of January one thousand eight hundred and one, all prohibitions and bounties on the exports of articles, the growth, produce, or manufacture of either country, to the other, shall cease.... That all articles, the growth, produce, or manufacture of either country... shall from thenceforth be imported into each country from the other, free from duty... (Act of Union, 1st August 1800; available at www.statutelaw.gov.uk)

Among ethnic identities living in Ireland, 90% are Irish people whereas in England the ratio of White British is 81.9%. About 700,000 Irish born people live in England. England's political relationships have never been healthy with Ireland. During the twentieth century in Northern Ireland, ethnonational conflict ignited ethnic fabric of both the nations. This conflict was primarily the political one, but later on it turned to be sectarian. Ireland's 84% population is Roman Catholic whereas England's 71.8% population is Protestant. Keeping in view the ethnic fragmentation of Anglo-Irish society, swearwords both in written and spoken expressions, show the disgrace, disrespect, and hatred against the opponent society, culture, and nationality. This is what BNC data show that the most swearwords have been used against Irish nationals. This is the highest ratio against Irish in BNC both in written and spoken language, respectively.

On the contrary, we do not find swearword in BNC for African Americans (Porch Monkey) and Jews (Heeb). One of the impressions for the absence of these swearwords may be on the credibility and representativeness of BNC. The political and social status of these two cultural and national identities has manifold cultural imprints on British society.

In BNC spoken corpus, the 2nd highest number of swearwords is against Pakistani nationals after Irish. BNC is a representative data from 1980 to 1993. Second highest number of swearwords against Pakistan reflects the socio-political relationship between both the nations. During the Zia-ul-Haq regime, many a sociopolitical changes took place in Pakistani society. On May 26, 1980 for the first time in Pakistan Federal Sharia Courts were established. In 1981, Interest-free banking was introduced, and on March 11, 1983 a nuclear test was carried out by Pakistan without announcing. Being a former colony of Great Britain such changes in indigenous policies and political narratives might not be expected from British, so they started to reflect their hatred particularly against Pakistani diaspora in spoken expression. This is evident from the BNC data that swearwords accommodate national interests of a nation.

6. Conclusion

After presenting the findings in detail, this section of the study offers a brief conclusion drawn from the study. Language as a tool defines the cultural and sociopolitical narratives in written (literary, journalistic, academic, etc.) and spoken text. The use of specific swearwords may provide evidence to reach the psychological and cultural changes that take place in a society. This study has provided us database evidence to deconstruct language and

society at the same time. The research questions posed in the introductory section have been addressed on the basis of findings from the BNC. To sum up, BNC has provided evidence about the selected swearwords that have been used by native speakers to channelize their invasive variants. Speakers, particularly in spoken variety of language, have been seen using highly aggressive tokens against some nationalities. This study also reveals that a number of nationality-based swearwords by native speakers reflect a psychological superiority complex and hatred against opposing nationalities. This has also been noticed that the conflicts of England with its neighboring nations have been reflected in written and spoken expression by the English. In addition, the paradigm shift in political and cultural narratives has also been reflected in written and spoken expressions in BNC. The use of swearwords by a particular society needs to be psychoanalysed in later studies.

References

- Anderson, L.G., & Trudgill, P. (1990). *Bad language*. England: Penguin Books.
- Arsana, M., Putra, I. N. A. J., & Ramendra, D. P. (2017). An analysis of balinese swear words among the northern balinese teenagers in seririt district in their daily conversation. *Jurnal Pendidikan Bahasa Inggris*, 3(3).
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language Structure and use*. Cambridge, UK: Cambridge University Press.
- British National Corpus (<https://corpus.byu.edu/bnc>).
- Bruin, M. D. (2018). *Translating swear words: a fucking difficult task. An analysis of Sam Garrett's American English translation of Turks Fruit* (Master's thesis).
- Chen, C. (2004). On the Hong Kong Chinese Subtitling of English Swearwords. *Meta: Journal des traducteurs/Meta: Translators' Journal*, 49(1), 135-147.
- Christie, C. (2013). The relevance of taboo language: An analysis of the indexical values of swearwords. *Journal of Pragmatics*, 58, 152-169.
- Darma, R. S., Wennyta, W., & Fitri, N. (2017). Swearing Words in “Celebrity Read Mean Tweets” in Jimmy Kimmel Late Night Show: A Sociolinguistic Study. *JELT: Journal of English Language Teaching*, 1(1).
- Dewaele, J. M. (2015). British ‘Bollocks’ versus American ‘Jerk’: Do native British English speakers swear more—or differently—compared to American English speakers? *Applied Linguistics Review*, 6(3), 309-339.
- Dewaele, J. M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of multilingual and multicultural development*, 25(2-3), 204-222.
- Fägersten, K. B. (2014). The use of English swear words in Swedish media.
- Fägersten, K. B. (2012). *Who’s swearing now? The social aspects of conversational swearing*. Cambridge Scholars Publishing.
- Fras, J. (2012). The ideology of swearwords in Slovenia. *Language & Communication*, 32(3), 229-239.
- Gauthier, M. (2012). Profanity and gender: A diachronic analysis of men's and women's use and perception of swear words. *Profanity and Gender: A Diachronic Analysis of Men's and Women's Use and Perception of Swear Words*.

- Goddard, C. (2015). "Swear words" and "curse words" in Australian (and American) English. At the crossroads of pragmatics, semantics and sociolinguistics. *Intercultural Pragmatics*, 12(2), 189-218.
- Gries, S. T. (2009). What is corpus linguistics? *Language and linguistics compass*, 3(5), 1225-1241.
- Gries, S. T. (2009). *Quantitative corpus linguistics with R: A practical introduction*. New York, NY: Routledge.
- Güvendir, E. (2015). Why are males inclined to use strong swear words more than females? An evolutionary explanation based on male intergroup aggressiveness. *Language Sciences*, 50, 133-139.
- Hjort, M. (2009). Swearwords in subtitles: A balancing act. *TRAlinea, Special Issue: The Translation of Dialects in Multimedia (2009)*.
- Harismayanti, I., Budasi, I. G., Lin, D. A., & Ramendra, D. P. (2017). An Analysis of Swear Words Used by the Teenagers in Singaraja. *Jurnal Pendidikan Bahasa Inggris undiksha*, 4(2).
- Jay, T. (1992). *Cursing in America: A Psycholinguistic Study of Dirty Language in the Courts, in the Movies, in the Schoolyards, and on the Streets*. Philadelphia: John Benjamins.
- Khoirunnisa, F. R., & Wijayanto, A. (2017). *The Use Of Swear Words In Deadpool Movie: A Sociopragmatic Study* (Doctoral dissertation, Universitas Muhammadiyah Surakarta).
- Kirk, C. (2013). The most popular swear words on Facebook. *Slate.com*.
- Kurniawan, K. E. (2015). *The analysis of taboo words and swear words in george carlin's monologue seven words you can't say on television* (Doctoral dissertation, SANATA DHARMA UNIVERSITY).
- Kwon, K. H., & Cho, D. (2016). Reimagining swearing effects online: the use of swearwords in news audience commenting. *AoIR Selected Papers of Internet Research*, 5.
- Leech, G., & Rayson, P. (2014). *Word frequencies in written and spoken English: Based on the British National Corpus*. Routledge.
- Ljung, M. (2010). *Swearing: A cross-cultural linguistic study*. Springer.
- Ljung, M. (2006). *Svordomsboken*. Uddevalla: MediaPrint.
- Marom, N. (2014). In Case of a Curse: *The Translation of Swearwords in Hebrew Subtitles of American Films and Television Series*. University of Haifa, Faculty of the Humanities, Department of English Language and Literature.
- McEnery, T. (2004). *Swearing in English: Bad language, purity and power from 1586 to the present*. Routledge.
- McEnery, A., & Xiao, Z. (2004). Swearing in modern British English: the case of fuck in the BNC. *Language and Literature*, 13(3), 235-268.
- Montagu, A. (1967). *The anatomy of swearing*. University of Pennsylvania press.
- Moore, R. L. (2012). On swearwords and slang. *American speech*, 87(2), 170-189.
- Praschinger, A., Pomikal, C., & Stieger, S. (2011). May I curse a referee? Swear words and consequences. *Journal of sports science & medicine*, 10(2), 341.
- Popușoi, S. A., Havârneanu, G. M., & Havârneanu, C. E. (2018). "Get the f##* k out of my way!" Exploring the cathartic effect of swearwords in coping with driving anger. *Transportation Research Part F: Traffic Psychology and Behaviour*, 56, 215-226.

- Rullyanti, M. (2017). Swearwords used by Bengkuluese: A Case Study of Taboo Expressions. *Premise: Journal of English Education*, 3(1), 18-27.
- Stenström, A. B. (2017). *Swearing in English and Spanish teenage talk. Advances in Swearing Research: New languages and new contexts*, 282, 157.
- Suyanto, A. (2010). Javanese swear words in a boarding house: The case of five Soegijapranata Catholic university students. *Unpublished thesis. Soegijapranata Catholic University, Semarang*.
- Thorne, T. (2014). *Dictionary of contemporary slang*. Bloomsbury Publishing.
- Zenner, E., & Ruetten, T. (2018). The borrowability of English swearwords in Dutch: a variationist approach.