

Review

Urdu in anglicized world: A corpus based study

Muhammad Asghar, Zobina* and Mahmood, Muhammad Asim

Government College University Faisalabad, Pakistan.

Accepted 24 April, 2013

This paper aims to find out the variation in Urdu language due to language contact. It studies the use of English vocabulary in Urdu written language. The nature of the research is quantitative as well as qualitative as it examines the grammatical and domain-wise categories of English words used in Urdu and tries to determine the social factors that are the cause of this increase. It counts and calculates the frequency and the percentages of English words to examine the diachronic changes in the language. The researchers have taken the corpus of 30,000 words in total which includes 15,000 words of Urdu magazine published from 1995 to 1997 and the similar corpus of year 2011. The corpus is minutely observed by the researchers and English words are extracted and tested from English to Urdu dictionaries in order to see whether or not English vocabulary is increasing due to (non)availability of their Urdu substitutes. The results reveal that Urdu language is now absorbing English words more than a decade ago. The results also brought forth an interesting fact that the magazines that are promoters of Urdu language are conscious to restrain the use of English. However, English words frequently used by Urdu users are showing up in the written language and are used at the expense of their Urdu correlates.

Key words: Code-mixing, code-switching, variation in Urdu.

INTRODUCTION

Languages are living organisms. They take birth, grow and die with the passage of time. Languages influence each other when they come into contact. This contact "is initiated by the spread of languages of power and prestige via conquest and colonization" (Mesthrie, 2009) and causes code-mixing, code-switching and lexical borrowing.

Urdu language evolved mainly from Persian and Arabic almost two centuries ago and "has a long history of amalgamation with other languages" (Sarwet, 2006). Since then, it has been expanding itself through accepting words from different languages with which its users come into contact. Colonizing forces of Britain made English and Urdu co-exist since the 19th century. The English language expanded itself in the new scenario. With the exclusion of the English from the subcontinent in the

middle of 20th century; however, the flow of Urdu into English decreased but the two languages continued sustaining in the new born state, that is, Pakistan. As "the social context largely affects the linguistics choices and preferences" (Sarwet, 2006), the social factor behind this linguistic phenomenon was that the Pakistanis who wanted Urdu as their national language have been surviving through the English language for the last hundred years. They could not draw a neat line between the uses of two languages and thus both languages were being used simultaneously. In the beginning, the influence of English phrases and words remained limited only to the spoken Urdu language because new state idealized Urdu as its state language. Literacy was promoted through national language but English being the language of the recent x-rulers was prestigious in

*Corresponding author. E-mail: zobinaasghar@yahoo.com. Tel: (+92)03428760876.

social setup and the elite class was observed to be borrowing English words in their speech.

In recent years, with the advent and spread of modern technological devices and globalization, this trend has trickled down to the common masses and has extended itself from spoken Urdu to written text. Even the books and magazines that are known as promoters of Urdu language around the world are unable to resist this flux.

Any language is thought to be growing if it expands its tether by accepting the words and concepts that it lacks itself, but a threat to its status is felt if the borrowed words damage its original structure or are spoken at the expense of the first language.

Aims of the study

The present study, at first, aims to investigate the assumption that English vocabulary is increasing in Urdu magazines. Secondly, it studies which grammatical category is borrowed more. It also investigates which field of life is being affected more by this code-switching and whether the English vocabulary has intruded unpleasantly to Urdu language or is playing a positive role in its expansion and growth. It also tries to locate the reasons and draws conclusions accordingly.

LITERATURE REVIEW

Code-switching, code-mixing and lexical borrowing are sometimes confused with each other. Before moving ahead it would be better to define these terms and the way they are used in this study. The term code mixing refers to mixing of two or more languages within a sentence while the term code switching refers to mixing of two or more languages at the clausal level in a discourse in a fully grammatical way (Poplack, 1980). Bhatia and Ritchie (2004) defined code-mixing as the mixing of various linguistic units (morphemes, words, modifiers, phrases, clauses and sentences) primarily from two participating grammatical systems within a sentence. Bhatt (1997) has used these terms interchangeably and in the present study no such distinction has been observed, as "code-switching within a sentence refers to code-mixing and borrowing" (Sarwet, 2006).

The present study will take intra-sentential switching as code-mixing and in the same sense it will use the term lexical borrowing.

English is now a lingua-franca and is used in almost every part of the world for in and out-group communication. It is excessively used in the outer-circle countries where it is learnt and taught as Second Language and where the know-how of English is associated with economic and social prosperity. Researchers have kept themselves busy in the investigation of reasons of code mixing and come up with various sociolinguistic and

psycholinguistic factors that contribute to code mixing and code switching. One such factor is bilingualism and language contact that results in the lexical borrowing and mixture of two languages in the bilinguals (Ansre, 1971; Bamgbose, 1971; Cheng and Butler, 1989). Other causes include modernization, westernization, efficiency, professionalism and social advancement (Kachru, 1989; Kamwangamalu, 1989). According to these scholars, some of the functions of code switching and code mixing are intra group identity (Gumperz, 1982), poetic creativity (Kachru, 1989) and the expression of modernization (Kamwangamalu, 1989).

Stanlaw (1987) pointed out that "one important reason for the use of English loan words is that they provide linguistic tools that individuals can use in personal and highly creative ways". However, Takashi (1997) did not totally agree with the idea and stated that "it seems inadequate to attribute the high number of English elements which do not fill lexical gaps to the loan words with modern connection alone". Takashi concludes that the innovative use of English catches the attention of the audience and it may motivate the combination of loanwords. Also, Piller (2003) suggests "foreign language... impedes the automatic processing and thereby arrests the attention of recipients for a longer time span". Some researchers go for other reasons of code-mixing with English. According to them, English is used as an association with success, internationalization, sophistication, future orientation and fun (Alm, 2003; Kelly-homes, 2005; Piller, 2003). English has been used as attention-getting device and is connected with advancement and modernity.

Nishimura (1997) investigates code-switching between Japanese and English among the second generation Japanese. In this study, the author identifies three types of bilingual speech: first, the basically Japanese variety with lexical English lexical items; second, the basically English variety with English as the base language; and third, the mixed variety with no base language identifiable. The study reveals that the same individual uses different types of code-mixing and the reasons for code switching vary.

Various researchers have also investigated the phenomenon of code-mixing. Shin (2002) explores the characteristics of code-mixing among Korean-English bilingual children. His main focus remains on the distinction between code switching and borrowing. Hussein (1999) examines the attitude of Arab teenage students towards code switching and code mixing. Also, he investigates the reasons for such linguistic behavior. The results of the study revealed that both positive and negative attitudes are found among the students about code mixing. It also explores that the most important reason for code switching is the lack of Arabic equivalents for English terms. Arnfast and Jorgensen (2003) consider code switching as an increasingly sophisticated and complicated language skill during second language learning and

Table 1. Overall findings.

Year	Total words	English words	age%
1997	15031	117	.8
2011	15200	185	1.2

explore the use of code-switching in the learners of Danish. Leung (2010) studies code mixing in print advertisement in Hong Kong. The study reveals that the use of English words in print advertisement has been associated with modernity and internationalization. In the present work, the code mixing of English and Urdu is observed in the written text of Urdu. The study observes the code-mixing over time, that is, from 1995/97 to 2011. The researchers have tried to find the results and implications of this code mixing to bring this phenomenon to the conscious level of the Urdu speakers.

In Pakistani scenario, Sarwet (2006) studied the phenomenon of code-mixing between Urdu and English in media language and noticed that through this phenomenon, English vocabulary is permeating into social domains of the Urdu speaker. For the purpose, she utilized the word categories devised by Kachru (1978), and which he has used to study the code-mixing between Hind and Urdu languages.

METHODOLOGY

The researchers have chosen a renowned Urdu magazine "*Phool*" (Pakistan) because it is read in more than fifty countries as Urdu magazine. The corpus is divided into two segments; first, the magazines published between 1995-1997, and second, 2011. Ten articles (fiction) of almost equal length have been randomly selected from each segment. English words have been extracted and grouped into two main categories; grammatical and domain-wise categories, using the content analysis approach. These two categories have been further sub-divided to determine which grammatical category and domain of life is more affected by the phenomenon under study. It is important to mention here that these sub-categories are data-driven and not pre-set or pre-defined. This classification highlights which field of life is more influenced by code-switching and through which grammatical category. For the purpose, the frequencies are counted and are presented in percentages.

DATA PRESENTATION AND DISCUSSION

The study of the data reveals that English vocabulary has increased in frequency with the passing years. The results show that in first segment of the articles (1995-1997), the total English words are 117 out of total 15031 words, while in second segment this number reaches to 185 among 15200. The calculation shows that this frequency amounts to .8 words per hundred words in first segment and 1.2% in the second. The increase is observed as .4% from 1997 to 2011 (Table 1).

At first glance, this increase seems to be insignificant

Table 2. Comparison of grammatical categories.

Year	1995-1997	2011
Total English words	117	185
Nouns	112	172
Adjectives	4	9
Verb	1	4
Adverb	1	-

but is important because it occurred in an Urdu magazine whose editor, when consulted, said that they constantly edit the English words from the writings and translate them into Urdu before publication. This is because the magazine works on an ideology of promoting Pakistan's official language Urdu, and culture around the world. It is important to mention here that this magazine is read in more than fifty countries as an Urdu magazine.

On categorizing the English word frequencies grammatically, the fact sprang up that nouns are the most frequently borrowed. Among the total words of 117 of the first segment, there are 112 nouns, only three adjectives, 1 adverb and one verb; and in second segment among 185 frequencies, there are 172 nouns, 9 adjectives and only 4 verbs (Table 2).

Grammatical categories

Nouns are borrowed from other languages when people come into contact with new things and do not have proper vocabulary. This is not always the matter of unavailability of words in first language; there is another reason for borrowing words from other language and using them at the expense of the speakers' /writers' own language, and that is prestige. The results of the present study correspond with these two reasons for word borrowing.

Most of the nouns are related to either modern technology or modern style of life; these nouns do not have proper Urdu alternative for example: camera, T.V Lounge, Cinema, Mobile phone, biscuit, and sandwich. The other category is of the nouns that have Urdu alternatives, but due to less use they seem less familiar even to the Urdu speakers. These nouns are lawn (*Baghicha*), warning (*ta(m)bih*), road (*sarrak*), flight (*parwaz*), airport (*hawai adda*) to mention but a few.

The Use of English vocabulary is considered to be the sign of an educated person, so the above English words among Urdu are never discouraged.

Names of degrees and certificates do not have their Urdu alternatives (Table 3); example, M.A, (Master of Arts), Inter (intermediate) etc. Many other English words are frequently used at the expense of their Urdu alternatives; example, class fellow (*hum jamaat*), class (*jamaat*), bag (*basta*), paper (*prcha*), homework (*ghar ka kam*), pen (*qalam*), break (*waqfa*), etc.

Table 3. Comparison according to the availability of Urdu substitutes.

Words	Words with Urdu substitutes	Words without Urdu substitute
Frequency	213	89
Percentage	70.5	29.5

Table 4. Comparison of domain-wise categories.

Domain-wise categories	Frequency-1997	Frequency-2011
Profession	12	19
Relations	4	1
School/education	35	47
Documents	1	6
Kitchen/eatables	9	4
Dress/accessories	7	7
Science/discovery/invention	5	16
Transport/vehicle	5	10
Misc	11	30
Law and order	1	8
Place	16	18
State/behavior	2	10
Entertainment	9	9
Total	117	185

Domain-wise categories

All English words extracted during this study have also been categorized in another way: they are listed according to their field, that is, they are categorized domain-wise. These categories are data driven and were not pre-specified. These categories are; profession, relationships, school/education, documents, kitchen/eatables, dress/accessories, science/discovery, transport/vehicles, law and order, place, state/behavior, entertainment and miscellaneous. This categorization showed that the most influenced field of life is school and education.

Table 4 shows the frequency occurrence according to the domains of social life.

In all domain-wise categories, the occurrences of English words have increased or remained the same; for instance: profession, from 12 to 19; documents, from 1 to 6; school/education, from 35 to 47, except relationship and kitchen/eatables (Appendix). In Pakistani middle class society, a trend related to the relationships has been observed that many people prefer to teach their children to use Pakistani cultural words for relations contrary to the trend previously grown. For example, now people teach their children to call their aunts and uncle as '*Khala*' (Mother's sister), '*Phuphu*' (Father's sister), '*Chachu*' (Father's brother), '*Mamu*' (Mother's brother) and the previous generation's '*Mama*', '*Papa*' are being replaced with '*Mama*' and '*Baba*'. So this decrease in

English vocabulary in the segment of 2011 is seen in this perspective; however, the decrease in English vocabulary related to eatables is an exception because this is opposite to the prevalent situation.

Conclusion

The results of the study reveal that English vocabulary is increasing in Urdu magazine with the passage of time. This increase is triggered both by the new inventions and terminologies used to move in a modern style of living. This conclusion is gathered from the fact that the English vocabulary is used both for new concepts as well as existing concepts in Urdu language. In written language, this code-mixing is observed at word level only, thus it is not marring the basic structure of Urdu language. "Nouns" are the most borrowed grammatical category, while in domain-wise categories; "school and Education" is influenced the most by code-switching practice. This borrowing is not solely due to the non-availability of Urdu correlates rather there are other issues involved, it may be prestige, social influence, phonological ease but this is the topic of a separate research which could study the attitudes of people and explore the reasons for using English words in Urdu. However, the findings of the present study showed that Urdu is a flexible language and it is openly absorbing foreign vocabulary without

facing any distortion in its structure.

ACKNOWLEDGEMENT

My special thanks to Zahida Hussain who offered her unconditioned cooperation during this study

REFERENCES

- Alm CO (2003). "English in the Ecuadorian Commercial Context", *World Englishes* 22(2):143-158.
- Ansre G (1971). *The Influence of English on West African Languages*. In: Spencer J (ed.), *The English Language in West Africa*. Ibadan: Ibadan University Press.
- Bamgbose A (1971). *The English Language in Nigeria*. In: Spencer J (ed.), *The English Language in West Africa*. Ibadan: Ibadan University Press.
- Bhatia TK, Ritchie WC (2004). "Social and Psychological Factors in Language Mixing", In: Bhatia TK and Ritchie WC (Eds.), *The handbook of bilingualism* Blackwell. pp.336-352.
- Bhatt R (1997). "Code-Switching, Constraints, and Optimal Grammars", *Lingua* 102, pp. 223-251. Publishing Limited, Malden.
- Cheng L, Butler K (1989). *Code-switching: A Natural Phenomenon vs. Language "Deficiency"*. *World Englishes* 8(3).
- Gumperz JJ (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- Kachru BB (1978). *Towards structuring code mixing: An Indian perspective*. Kachru B, Sridhar SN (Eds.). *International Journal of Sociology of Language-Aspects of Sociolinguistics in South Asia*. (16). New York: Mouton Publishers
- Kachru Y (1989). *Code-mixing, Style Repertoire and Language Variation: English in Hindu Poetic Creativity*. *World Englishes* 8(3).
- Kamwangamalu N (1989). *Code-mixing and Modernisation*. *World Englishes* 8(3).
- Kelly-Homes H (2005). *Advertising as Multilingual Communication*, Palgrave Macmillan, New York.
- Leung C (2010). *Code-Mixing in Print Advertisement and its Cultural Implications in Hong Kong*. *Eur. J. Soc. Sci.* 12(3):417-429.
- Mesthrie RJS (2009). *Introducing Sociolinguistics*. Edinburgh: Edinburgh University Press.
- Piller I (2003). "Advertising as a Site of Language Contact", *Annu. Rev. Appl. Linguist.* 23:170-183.
- Poplack S (1980). *Sometimes I'll start a sentence in English y termino en espanol*. *Linguistics* 18:518-618.
- Stanlaw J (1987). "Japanese and English: Borrowing and Contact", *World Englishes* 6(2):93-109.
- Takashi K (1997). "Japanese Bilingual Brand Names", *English Today* 52:12-16.

Appendix. Domain-wise categories: 1995-97.

Profession/ employment	Relations	School/35 education	Documents1	Kitchen/ Eatables9	Dress/ Accessories7	Science/ Discovery Invention5	Transport/ Vehicle5	Misc11	Law and order1	Place 16	State/2 behaviour	Entertainment9
Hunter Senior officer Policeman Teacher Driver Doctor Police Heroine Business Compere engineer	Uncle Dad Mummy cousins	Class fellow Seat Class Period Sir (teacher) Bag Paper Tuition School College Homework Position Lecture School bag Pen Tuition Break Bag School 1 st year Progress report Top Students Teachers Stage Study School Test Chemistry Madam Notes Position Board Principal inter	ticket	Thermos Biscuit Ice cream Biscuit Cake Chocolate Crockery Plate Dinner	Dresses Suit Jacket Ready made(dress) Jewellery Purse handbag	Black hole Dinosaur telecast camera T.V	Engine Start Car Van taxi	Cylinder Shift Set Design design bed Christmas poster training field amnesty Super hit	FIR	Park Workshop Drawing room colony General store airport gallery hall T.V lounge Ground Cinema Drawing room Stage ground T.V lounge gallery	late Boycott	Drama serial Function Film Record film Hockey Ball Video games Games

Appendix. Domain-wise categories: 2011.

Profession/ employments 19	Relations 1	School/ education	47	Documents6	Kitchen/ Eatables4	Dress/ Accessories7	Science/ Discovery Invention16	Transport/ Vehicle/roads10	Misc30	Law and order8	Place 18	State/behaviour10	Entertainment9
Bomb disposal squad Commando FC Special services group Agency Principal Clerk Inspector professor teacher teacher doctor secretary police driver clerk doctor government officer teachers	mummy	Fees Class Desk Library Period Book College gate M.A Color Drawing Pencil Copy Chart Column Result sheet Science Unit Assignment Result School2 Uniform First year Sir2 Class	Lunchbox Group Test Library Fees college second year B.A Inter Matric School Certificate Lecture Break Groupfellow Desk Classroom Bag School Bench Assembly	Report Form Post card Note Money order form Certificate	dining table biscuit tray sandwich	Jacket Pant Shirt Jewellery Hanger coat knicker	Drone Missile Rocket Bomb TV2 channel Trigger Science Clock Room cooler Larva Platelets Dengue Biology Mobile phone	One way GT road Motorcycle tyre Tram Train Railway Mont pleasant road Car flight	Alert Duty Front line Furlong Target Training Stop Imported Model Order Drip stand plan type second hanger kick entry out of focus position drum mission marble line clinic let injection private interview address world tour	Security firing siron Blast Rally Service revolver Accident Police	cinema centre market ground lounge farmhouse drawing room market ground bathroom classroom swimming pool lawn railway station maingate counter gate hall	Disturb2 Sorry Warning Late Mood Idea Thank you Off mood Fool	Cinema party Match wicket Bat Fielding Football Cricket funfair