
A Corpus Based Quantitative Survey of the Persian and Arabic Elements in the Basic Vocabulary of Urdu Language

_____ María Isabel Maldonado García
=====

Abstract

This article is one of a series of articles in regards to the density of the etymological elements which compose Urdu language. In this particular article the exact density of Persian and Arabic elements will be obtained. Urdu is an Indo-European language which derived from Sanskrit into the Prakrit languages and at a later stage was influenced by Persian and Arabic languages, due to both, cultural and religious' influences. Hindi and Urdu were, in fact, the same language which separated due to the split of the script (digraphia). The Muslims started using a Persian script called Nastaliq, while the Hindus kept on using the Devanagari script. After the split of the language Urdu received numerous loans and grammar structures from Persian and Arabic. The research is noteworthy since it is almost certainly the first corpus-based survey of the use of Persian and Arabic derived terms in Urdu language. The research is conducted utilizing one thousand words of the basic vocabulary of Urdu as the corpus and analyzing them from an etymological point of view. It is conducted through a quantitative, descriptive and analytical

methodology, which will reveal the precise percentage of Persian and Arabic vocabulary present in Urdu language.

Historical Context

The history of Urdu language is without any doubt very peculiar and different from that of other languages. Many socio-cultural issues can change the demographics of a language as well as its inherent characteristics. The large majority of linguists agree that Urdu language derived from Sanskrit¹ into the languages known as the Prakrits. One of those languages was Khari Boli, a language spoken in the north of India from where Hindi and consequently, Urdu and Hindi languages derived.²

In this sense, Urdu and Hindi have suffered a process of evolution as well as reforms until its modern and present state.³ The reforms of Urdu language include but are not limited, in a chronological order, to extensive lexical and grammar borrowing from Persian language during the pre-Islamic period, extensive borrowing from Arabic with the advent of Islam from 712 A.D., at the same time that the Persian influence continued. These were mainly due to politico-religious reasons which initiated the split of Urdu and Hindi, although literary movements generated an enormous modifying influence. Other types of influence were of an artificial nature (not due to a natural evolutionary progression of the language), for example those which occurred during the time of the Mughals⁴ when the purging of the Sanskrit elements from the language occurred, substituting them with elements from Persian and Arabic, the Iranian imagery and metaphors replaced the Indian, and words of local dialects were also replaced with Persian and Arabic elements.⁵

Extensive types of reforms took place during the pre-Muslim period, the Muslim rule, the British period, the pre-partition period as well as the post partition period. All of them seem to be equally important if considered through the light of linguistic development and an evolutionary point of view. Nevertheless, the event which really marked the difference between Urdu and Hindi languages was the split of the script.

The Script

Urdu and Hindi languages were in fact one and the same. The difference between them has its origin in socio-linguistics and cultural factors related to a digraphia phenomenon by which Urdu language is written using the Nastaliq script and Hindi language is written using Devanagari script. King bifurcates the languages on the basis of the script:

Digraphia is regularly an outer and visible sign of ethnic or religious hatred. Script tolerance, alas, is no more common than tolerance itself. In this too Hindi-Urdu is lamentably all too typical. People have died in India for the Devanagari script of Hindi or the Perso-Arabic script of Urdu. It is rare, except for scholars, for Hindi speakers to learn to read Urdu script or for Urdu speakers to learn to read Devanagari.⁶

This extreme case of digraphia is significant in the case of Urdu for several reasons. In India two scripts were mainly utilized, the Kharoshthi script, possibly derived from Aramaic,⁷ although most of the scripts utilized in India seem have derived from the Brahmi script.⁸ During the Persian rule, Kharoshthi was mainly utilized.

At the present time Hindi utilizes the Devanagari script (from Brahmi), written from left to right, connecting words and letters by an overhead horizontal line. However, Urdu employs the Nastaliq script which was created approximately around the 1400's⁹ utilizing the *Nashk* (or *Nashki*) of Arabic and the outdated Persian *tal'iq*. Its conceiver seems to have been the renowned calligrapher Mir Ali Tabrizi¹⁰ (1360-1420). This seems to be the turning point for the appearance of the two languages as individual entities. Nevertheless, the language had received extensive influence from Persian and Arabic languages even before that.

The visual differences between the scripts that are nowadays utilized for both languages are noteworthy. They can be observed underneath:

Devanagari script:	एक छोटी सी मेज
Nastaliq script:	ایک چھوٹی سی میز

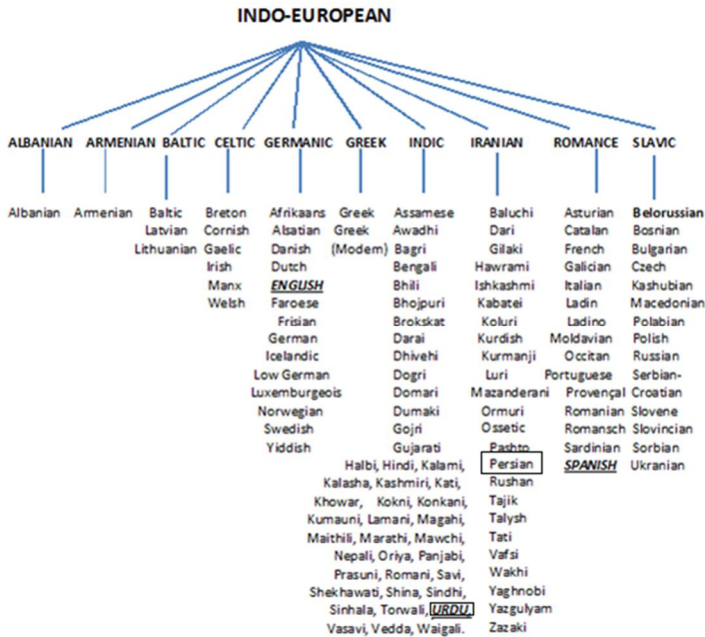
In this sense, it is important to note that an Arabic script does not suit Urdu or any Indo-European language, since Arabic is a Semitic language which does not need vowels to be read and Indo-European languages are highly dependent on a vowel system. Nevertheless, in the past, cases have been observed where a language would utilize a script which belongs to a language of a different family altogether, such is the case of Turkish language,¹¹ which, due to a political language reform by the ruler of the time Ataturk, utilizes a Latin script effectively, although it does not belong to the Indo-European family of languages. After the split of the script Urdu continued receiving loanwords from Arabic, Persian and to a minor extent Punjabi and Turkish, while Hindi, continued borrowing important terminology from Sanskrit language.

India's linguistic history spans more than three thousand years. Historically, Sanskrit had developed into the Prakrit languages such as Marathi, Gujarati, Oriya, Bengali and Khari Boli, a branch of Western Hindi and Hindvi from which Urdu as well as Hindi had developed.¹² Prakrit was later replaced by some of the dialects as early as the 11th century.¹³

It is a popular belief, supported by some scholars and philologists that Urdu is a mixture of languages such as Persian, Arabic and Turkish. It must be clarified that languages are alive and in continuous development. In this sense, as stated before Sanskrit developed into the Prakrits, which in turn developed into other dialects resulting in Urdu and Hindi. It was not a sudden and unnatural mixture of languages which came into being due to the presence in the area of the Persian, Turks and Arabs. Rather, the existing language received external linguistic influences due to the socio-cultural events occurring in the area at a particular point in time.

Influence From Arabic and Persian in Urdu Language
 Urdu (and Hindi) as well as Persian are Indo-European Language as the following figure depicts:

Figure 1
 The genetic relationship between Urdu and Persian languages



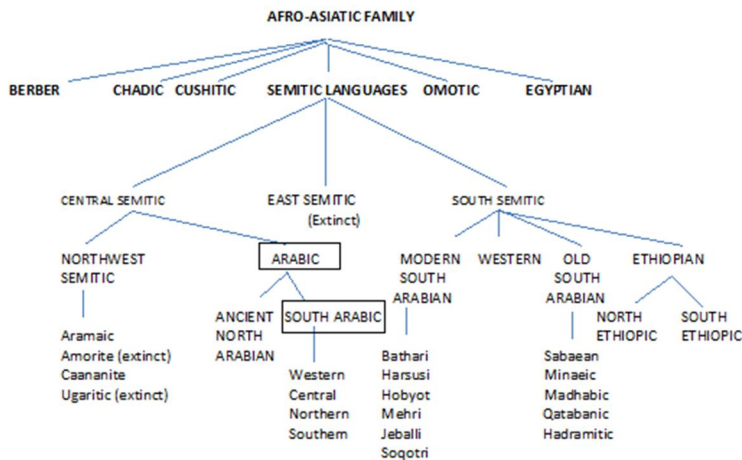
Source: Comparación del Léxico Básico del Español, el Inglés y el Urdu.¹⁴

In this sense they are related by a genetic linguistic link and belong to the same family of languages and different branches. Persian is an Indo-European language from the Iranian family of languages. It derives from Proto-Iranian language and it has undergone three stages or main periods, also called Old Persian, Middle Persian and New or Modern Persian. At every stage it utilized a different script. During the Old Persian period from circa 525-300 B.C. the script was a cuneiform script. In the Middle Persian period circa 300 B.C. -800 A.D. three scripts were utilized, they were the Pahlavi, Manichaean and the Avestan

script. During the Modern Persian period from 800 A.D. onwards Persian language utilized a Perso-Arabic script. Old Persian and Avestan are the oldest of the Iranian languages. Old Persian was the language of the people of the Achaemenid period.¹⁵ The language seems to have arrived to the Subcontinent with the Sassanian Aryan invasion, being the first of the Sassanids, Ardashir I (circa 230-250) the grandson of Sasan. The Sasanian period (circa 240-621) utilized Middle Persian language or Pahlavi. Persian dynasties kept their rule on certain areas of India, century after century which propagated the use of Persian language during the pre-Islamic period. In fact, the Persian culture has had a profound cultural impact on the Subcontinent¹⁶ and it seems to have been the language of the court.¹⁷ During this time Persian language kept on influencing the ancestors of Urdu language. Learning centers were established like the Nestorian learning center of Jundhishpur wherein translations of the Greek archaic works as well as works of other nations including Indian were being translated.¹⁸ By the 6th century and under Khosro I (circa 531-579) medical books were collected from India and translated from Sanskrit to Pahlavi, as well as the translations of other works of science were taking place from Greek into Persian.¹⁹

Arabic language in India was in fact, purely a foreign influence since it is not an Indo-European language and as such, not related to Urdu, neither Persian. Its Semitic nature was absolutely extraneous to Urdu. The following figure illustrates the Afro-Asiatic family of languages to which Semitic languages belong:

Figure 2
Afro-Asiatic Family of Languages



Source: Information for the figure taken from Ethnologue and The World Atlas of Language Structures²⁰

The beginning of the Arabic influence in Urdu language is very clear as it was marked by the arrival of the Muslims in the Subcontinent in 712 A.D. Arab traders had started their travels to India during this time which in turn produced Turko-Afghans invasions by the 10th century and by the 13th century the north of India almost in its majority was under Muslim rule.²¹

Tariq Rahman calls the process of Arabization of Urdu language “Muslimization of Urdu.”²² The expression actually fits the situation of the time to perfection as Iran had already become a Muslim state and Islam had been spreading in India. Important reformers of the time were Amir Khan (Mughal King Muhammad Shah’s minister), Amir Khusrau (1253-1325), Sirajuddin Ali Khan Arzu (1687-1756), Jān-e Janān (1701–1780), Mīr Dard, MirzāMuḥammad Rafī Sauda (1706–1781), Insha Ullah Khan Insha (1752/6-1817/8). In any case, an important factor in the new orientation towards Persian and Arabic was the Muslim reality itself which obliged to utilize terms not present in the

language, thus facilitating the borrowing in this sense, especially from Arabic.

However, measuring the individual exact impact of Persian and Arabic is a grueling enterprise and perhaps an impossible task. Nonetheless, Persian language seemingly has influenced Urdu longer and more intensely. This is due mainly to the fact that Persian language exercised a longer period of influence since it had a presential advantage over Arabic in India of approximately five hundred years; circa 240 to 1834/1837 when the British government discontinued the use of Persian language²³ and Urdu substituted it (with all the Persian influence already immersed in the language which provided an environment of continuity). Also due to the fact that Persian had been the language of the administration, courts and imperial rule, in fact, a language utilized for the purpose of communication and literature, not just of Islamic reality. Furthermore, Urdu maintained some of the grammar structures, lexicon as well as the script of Persian. On the other hand, Arabic is the language of the religious scriptures of Muslims Muslims on both sides of the border learn to read it from childhood. It was the source of extensive lexicon borrowing directed towards favoring an Islamic environment and creating an identity of the Muslims in India and Pakistan.

Measuring the Persian and Arabic Lexical Density in the Basic Vocabulary of Urdu Language

The objective of this paper is in fact, to measure the lexical density still remaining in the basic vocabulary of Urdu language which belongs to the Persian and Arab legacy. For this purpose the Persian and Arabic elements in the basic vocabulary of Urdu language will be identified, as well as measured. Then the basic vocabulary of Urdu language will be analyzed from an etymological point of view, in this way revealing the exact percentage of Persian and Arabic vocabulary present in Urdu's basic vocabulary.

Thus the basic vocabulary of Urdu Language is as follows:

صحیح ، سرگرمی ، سرگرم ، عمل ، پار ، ختم ، اوپر ، بارے میں ، ایک ، اکیلے ، تمام ، ہوا ، پہلے ، عمر ، پھر ، بعد ، ڈر ، مناسب ، شامل ، اپنانا ، جواب ، دوسرے ، غصہ ، ایک ، شمار ، ہوں ، ہمیشہ ، پہلے ہی ، ساتھ ، منظور ، سیب ، ظاہر ، کسیبھیوقت ، کچھہی ، کسی ، کوئی ، چیونٹی ، میں ، پوچھنا ، جیسا ، آرٹ ، پہنچ ، گرد ، فوج ، بازو ، علاقہ ، بین ، بینک ، گیند ، بستہ ، برا ، کمر ، بنیاد ، بے بی ، دور ، خزاں ، خالہ ، حملہ ، بستر ، شراب ، خوبصورت ، برداشت ، دانہ ، ہونا ، غسل ، ٹوکری ، بہترین ، علاوہ ، نیچے ، گھنٹی ، شروع ، پہلے ، برتاؤ ، سونے کاکمرا ، سیاہ ، کاٹ ، تھوڑا ، سالگرہ ، پیدائش ، پرندہ ، بڑا ، درمیان ، بہتر ، بڈی ، ابال ، جسم ، کشتی ، بورڈ ، نیلا ، پھونکنا ، خون ، بلاک ، خونہنا ، لڑکا ، ڈبہ ، کلٹورا ، نچلا ، بوتل ، دونوں ، قرض ، پیدا ، سرحد ، کتاب ، بھائی ، لانا ، روشن ، پل ، سانس ، ناشتہ ، ٹوٹنا ، روٹی ، بہادر ، شاخ ، خریدنا ، لیکن ، مصروف ، بس ، کاروبار ، جلا ، تعمیر ، برش ، بھورا ، احتیاط ، دیکھبھال ، کار ، ٹوپی ، مومبتی ، کرسکتا ، بلانا ، کیک ، نے ، کرسی ، بعض ، صدی ، مرکزی ، بلی ، کیس ، اٹھانا ، گاجر ، لاپروا ، انتخاب ، چاکلیٹ ، بچہ ، چکن ، پنیر ، سستا ، پیچھا ، تبدیل ، موقع ، چڑھنا ، واضح ، صاف ، چالاک ، جماعت ، شہر ، دائرہ ، انتخابکرنا ، سردی ، سکہ ، کوٹ ، کافی ، بند ، ابرالود ، بادل ، کپڑوں ، کپڑا ، گھڑی ، کمپیوٹر ، مکمل ، آنا ، موازنہ ، عام ، آرامدہ ، کنگھی ، رنگ ، جمع ، لاگت ، درست ، کونا ، مکئی ، تانبہ ، ٹھنڈا ، پکانا ، کنٹرول ، جاری ، شرط ، الماری ، کپ ، رونا ، پار ، تباہی ، ڈھانکنا ، راستہ ، ملک ، شمار ، شامل ، گہرا ، کم ، فیصلہ ، مردہ ، دن ، بیٹی ، اندھیرا ، خطرناک ، رقص ، کاٹنا ، سمت ، ڈنر ، مشکل ، مختلف ، مرنا ، ترقی ، تباہ ، ڈیسک ، انحصار ، ہرن ، کھینچنا ، نیچے ، دگنا ، نقطہ ، دروازہ ، کتا ، کرنا ، ڈش ، دریافت ، میلا ، کان ، ہر ، فرض ، مٹی ، بطخ ، خشک ، قطرہ ، چلانا ، پینا ، کپڑے ، خواب ، آتہ ، انڈہ ، اثر ، تعلیم ، کھانا ، آسان ، مشرق ، زمین ، کمانا ، ابتدائی ، بہت ، لطفاندوز ، دشمن ، آخر ، خالی ، اور ، ہاتھی ، بجلی ، کوئیبھی ، قطعی ، سب ، ہر ، ہمیشہ ، واقعہ ، شام ، بھی ، فرار ، داخلہ ، برابر ، داخلہونا ، مہنگی ، توقع ، ورزش ، پرجوش ، سوائے ، مثال ، امتحان ، سب ، مشہور ، خاندان ، گرنا ، جعلی ، ناکام ، حقیقت ، چہرہ ، آنکہ ، انتہائی ، وضاحت ، مؤنث ، محسوس ، کھلانا ، خوف ، غلطی ، تیز ، والد ، چربی ، فارم ، دور ، پہلے ، آگ ، ختم ، انگلی ، ٹھیک ، تلاش ، فلم ، بھرنے ، لڑنے ، چند ، بخار ، پرواز ، پھول ، آٹا ، فرش ، تیرنا ، ہموار ، پرچم ، ٹھیککرنا ، پانچ ، موزوں

غیرملکی، مجبور، کے لئے، فٹبال، پاؤں، بیوقوف، کھانا، تہکنا، منجمد، آزادی، مفت، چار، لومڑی، فارم، کانٹا، معاف، بھول، جنگل، مضحکہخیز، مزہ، بھراہوا، پھل، سامنے، سے، دوستانہ، دوست، تازہ، حاصل، شریفآدمی، عام، دروازہ، باغ، کھیل، مستقبل، مزید، فرنیچر، دادی، دادا، الوداع، اچھا، خدا، بکری، جانا، شیشہ، خوش، دینا، تحفہ، بال، بندوق، بڑھنا، گروپ، زمین، سرمی، سبز، عظیم، قبر، گھاس، رکھنا، نفرت، ٹوی، مشکل، خوش، ہونا، ہاتھ، ہتھوڑا، بال، نصف، مدد، آداب، اونچائی، جنت، بھاری، دل، سن، صحتمند، سر، اس، انکی، اس، پہاڑ، اونچا، چھپانا، اسکا، یہاں، اس، مرغی، مددگار، بوٹل، گرم، ہسپتال، گھوڑا، امید، گھر، چھٹی، چھید، پکڑنا، شوق، مارا، برف، میں، تکلیف، شوہر، جلدی، گھنٹہ، بھوکا، سو، کیسے، مکان، مدعو، لوہا، ایجاد، متعارف، میں، اندر، اضافہ، میں، اہم، اگر، خیال، صرف، چھلانگ، رس، شامل، نوکری، حبلی، اس، یہ، جزیرہ، بے، دستک، چاقو، گھنٹے، باورچیخانہ، بادشاہ، قسم، قتل، اہم، رکھنا، بعد، حال میں، دیر، آخری، بڑا، زمین، چراغ، خاتون، سیڑھی، معلوم، لمبائی، قرض، بائیں، ٹانگ، چھوڑدینا، سیکھنا، پتی، قیادت، سست، جیسے، روشنی، زندگی، جھوٹ، لائبریری، خط، اجازت، سبق، کم، کھو، نظر، طویل، تنہا، تالا، رہنا، تھوڑا، سننا، فہرست، ہونٹ، شیر، آدمی، نر، بنانا، اہم، مشین، قسمت، نیچے کرنا، کم، محبت، بہت، مطلب، کھانا، میں، شاید، بات، شادی، مارکیٹ، نشان، نقشہ، بہت، ملین، دودھ، درمیان، طریقہ، ذکر، رکن، ملاقات، دوا، گوشت، پیمائش، ماہ، بندر، پیسے، لمحہ، جدید، ماڈل، ملانا، غلطی، کھونا، منٹ، ذہن، موسیقی، بہت، حرکت، منہ، پہاڑ، ماں، زیادہ، صبح، مزید، چاند، گردن، تقریباً، قریب، نوعیت، قوم، تنگ، نام، میرے، ضروری، اگلا، اخبار، خیر، نئے، کبھی نہیں، جال، نہیہ، پڑوسی، سوئی، ضرورت، ناک، شمال، نہی، کونینہیں، شور، عظیم، نہیں، نو، رات، اچھا، کا، بحر، اعتراض، اطاعت، تعداد، اب، اطلاع، کچھبھی نہیں، نہیں، مخالف، کھول، صرف، ایک، پر، پرانے، تیل، اکثر، دفتر، پیشکش، درد، صفحہ، اپنے، اوپر، باہر، ہمارے، دیگر، حکم، نارنجی، یا، گزرنا، پارٹی، ساتھی، حصہ، پارک، والدین، کاغذ، بین، جوڑی، پینٹ، کامل، فی، مرچ، لوگ، پنسل، قلم، امن، ادا، راستہ، تھپکی، گزشتہ، سور، ٹکڑا، تصویر، چننا، پیانو، فوٹوگراف، پٹرول، شخص، مدت، خوش، پلینز، کھیلنا، پلیٹ، پلاسٹک، پودا، طیارہ، جگہ، گلابی، پن

مقام ، مقبول ، غریب ، پول ، شائستہ ، پولیس ، زہر ، نقطہ ، جیب ، بہت ، قیمت ، روکتھام ، خوبصورت ، پریس ، پیش ، طاقت ، ڈالنا ، آلو ، ممکن ، حفاظت ، مناسب ، وعدہ ، پیدا ، مسئلہ ، شاید ، انعام ، ذاتی ، جیل ، شہزادہ ، فوری ، سوال ، ملکہ ، ڈال ، دھکا ، طالبعلم ، سزا ، کھینچنا ، عوامی ، فراہم ، تیار ، پڑھنا ، پہنچنا ، بڑھانے ، برسات ، بارش ، ریڈیو ، قطعی ، خاموش ، مرمت ، کرایہ ، ہٹانا ، یاد ، لال ، ریکارڈ ، وصول ، واقعی ، اصلی ، امیر ، چاول ، واپس ، نتیجہ ، ریستوران ، آرام ، رپورٹ ، جواب ، دہرانا ، بدتمیز ، ربر ، گول ، کمرہ ، پتھر ، لوٹ ، سڑک ، بجنا ، رائٹ ، سواری ، اسی ، نمکین ، نمک ، محفوظ ، اداس ، جلدی ، بھاگنا ، حکمران ، حکمرانی ، دیکھنا ، دوسری ، سیٹ ، تلاش ، سائنس ، اسکول ، کہنا ، بچانا ، ریت ، سایہ ، رنگ ، جنس ، کئی ، سات ، خدمت ، سزا ، بھیجنا ، فروخت ، لگتا ، جہاز ، چمک ، شیلیف ، چادر ، بھیڑ ، صرف ، وہ ، تیز ، حصہ ، شکل ، ہلا ، بیمار ، ظاہر ، چیخنا ، کندھا ، مختصر ، دکان ، شوٹ ، جوتا ، قمیض ، ایک ، سادہ ، ملتے جلتے ، چاندی ، حماقت ، خاموشی ، اشارہ ، کنارہ ، سکرٹ ، جلد ، مہارت ، سائز ، چہ ، بیٹھا ، بہن ، ڈوبنا ، گانا ، بعدسے ، تمباکونوشی ، دھواں ، مسکرانا ، بو ، چھوٹا ، سست ، پھسلنا ، سونا ، آسمان ، بیٹا ، کبھی کبھی ، کچھ ، کسی ، کچھ ، نرم ، موزا ، صابن ، تو ، برف ، رفتار ، خصوصی ، بات ، خلا ، جنوب ، سوپ ، آواز ، معافکرنا ، جلدپی ، ستارہ ، کھڑا ، مہر ، مربع ، موسمیہار ، پھیل ، کھیل ، چمچ ، خرچ ، جادو ، پیٹ ، بھی ، اب ، قدم ، بھاپ ، چوری ، رہنا ، سٹیشن ، ریاست ، شروع ، ساخت ، مضبوط ، گلی ، عجیب ، کہانی ، طوفان ، ذخیرہ ، روکنے ، پتھر ، غیرمتوقع ، ایسے ، کامیاب ، مادہ ، مضمون ، پاگل ، مطالعہ ، طالبعلم ، حیرت ، یقین ، حمایت ، دھوپ ، سورج ، موسمگرما ، مناسب ، شکر ، پڑھانا ، چائے ، ٹیکسی ، ذائقہ ، اونچا ، بات ، لینا ، میز ، تلوار ، تیر ، میٹھی ، امتحان ، خوفناک ، ٹینس ، دس ، بتانا ، ٹیلیویژن ، ٹیلیفون ، آنسو ، ٹیم ، تیسری ، سوچتے ، چیز ، پتلی ، گاڑھا ، ان ، لہذا ، ادھر ، تو ، ان ، کہ ، سے ، ملکر ، پنچہ ، آج ، پر ، عنوان ، باندھنا ، صاف ، تین ، خطرہ ، اگرچہ ، اس ، ٹرام ، ٹرین ، ٹاؤن ، چھو ، ٹوٹل ، اوپر ، دانت ، آلہ ، بھی ، اجرات ، کل ، باری ، کوشش ، دو ، دوبار ، اعتبار ، سچ ، ٹرک ، مصیبت ، درخت ، سفر ، مفید ، استعمال ، اپ ، جبتک ، اکائی ، سمجھ ، نیچے ، چچا ، قسم ، چلنا ، جاگنا ، انتظار ، دورہ ، آواز ، گاؤں ، بہت ، سیزی ، اکثر ، حسبمعمول ، موسم ، پہننا ، ہم ، راستہ ، پانی ، دیکھنا ، ضائع ، دھونا ، گرم ، چاہنا ، کیا ، گیلی ، مغرب ، ٹھیکٹھاک ، خوشامدید ، عجیب ، وزن ، ہفتہ ، شادی ،

جنگلی ، بیوی ، وسیع ، کیوں ، جو ، سفید ، جبکہ ، جس ، جہاں ، جب ، پھپہا ، دانشمندانہ ، تار ، موسمسرما ، شراب ، کھڑکی ، ہوا ، حیت ، وصیت ، لکھنا ، بہتیرا ، فکر ، دنیا ، کام ، لفظ ، تعجب ، عورت ، بغیر ، ساتھ ، خواہش ، زیبرا ، صفر ، آپکا ، جوان ، آپ ، ابھیتک ، کل ، جی ، سال ، غلط ،

Once the etymology of all terms was extracted, the elements with etymology in Persian and Arabic were separated in different tables as follows²⁴:

Figure 3
Terms derived from Persian language

	English Meaning	Urdu term	Persian language
1	About	کے بارے میں	X
2	Across	پار	X
3	Active	سرگرم	X
4	activity	سرگرمی	X
5	Always	ہمیشہ	X
6	Am	ہوں	X
7	amount	شمار	X
8	another	دوسرے	X
9	arm	بازو	X
10	autumn	خزاں	X
11	base	بنیاد	X
12	back	برا	X
13	bag	بستہ، تھیلا	X
14	bean	دانہ، بیج	X
15	bed	بستر	X
16	below	نیچے	X
17	best	بہترین	X
18	better	بہتر	X
19	between	درمیان	X
20	bird	پرندوں	X
21	birth	پیدائش	X
22	birthday	سالگرہ	X

23	black	سیاه	X
24	bleed	خون بهنا	X
25	blood	خون	X
26	blue	نیلا	X
27	boat	کشتی	X
28	border	سرحد	X
29	born	پیدا	X
30	bottom	نچلا	X
31	branch	شاخ	X
32	brave	بہادر	X
33	breakfast	ناشہ	X
34	bright	روشن	X
35	business	کاروبار	X
36	but	لیکن	X
37	buy	خریدنا	X
38	candle	موم بتی	X
39	careless	لاپرواہ	X
40	century	صدی	X
41	change	تبدیل	X
42	cheese	پنیر	X
43	chicken	چکن	X
44	child	بچہ	X
45	city	شہر	X
46	cloudy	ابر آلود	X
47	close	بند	X
48	cold	سردی	X
49	color	رنگ	X
50	Comfortable	آرام دہ	X
51	correct	درست	X
52	count	شمار	X
53	course	راستہ	X
54	crash	تباہی	X
55	cross	پار	X
56	dead	مردہ	X
57	decrease	کم	X

58	destroy	تباہ	X
59	discover	دریافت	X
60	door	دروازہ	X
61	down	نیچے	X
62	dream	خواب	X
63	dry	خشک	X
64	duck	بطخ	X
65	each	ہر	X
66	east	مشرق	X
67	enemy	دشمن	X
68	equal	برابر	X
69	escape	فرار ہونا	X
70	evening	شام	X
71	ever	ہمیشہ	X
72	every	ہر	X
73	excited	پرجوش	X
74	exercise	ورزش	X
75	fat	چربی	X
76	fast	تیز	X
77	find	تلاش	X
78	flag	پرچم	X
79	flat	ہموار	X
80	Fly	پرواز	X
81	four	چار	X
82	free	مفت	X
83	freedom	آزادی	X
84	fresh	تازہ	X
85	friend	دوست	X
86	friendly	دوستانہ	X
87	fun	مزہ	X
88	garden	باغ	X
89	gate	دروازہ	X
90	glad	خوشی	X
91	glass	شیشہ	X
92	god	خدا	X

93	green	سبز	X
94	grey	سرمئی	X
95	ground	زمین	X
96	gun	بندوق	X
97	happy	خوش	X
98	head	سر	X
99	hen	مرغی	X
100	hope	امید	X
101	hot	گرم	X
102	husband	شوهر	X
103	ice	برف	X
104	If	اگر	X
105	job	نوکری	X
106	king	بادشاه	X
107	knock	دستک	X
108	lamp	چراغ	X
109	land	زمین	X
110	late	دیر	X
111	lazy	سست	X
112	less	کم	X
113	life	زندگی	X
114	light	روشنی	X
115	lion	شیر	X
116	list	فہرست	X
117	lonely	تنہا	X
118	low	کم	X
119	lower	نیچے کرنا	X
120	mark	نشان زد	X
121	marry	شادی	X
122	may	شاید	X
123	measure	پیمائش	X
124	meat	گوشت	X
125	middle	درمیان	X
126	month	ماہ	X
127	name	نام	X

128	narrow	تنگ	X
129	neck	گردن	X
130	neither	نه به	X
131	noise	شور	X
132	nor	نه بهی	X
133	offer	پیشکش	X
134	orange	نارنجی	X
135	other	دیگر	X
136	pain	درد	X
137	path	راسته	X
138	pink	گلابی	X
139	place	جگہ	X
140	plane	طیارہ	X
141	pleased	خوش	X
142	poison	زہر	X
143	polite	شائستہ	X
144	present	پیش	X
145	pretty	خوبصورت	X
146	prince	شہزادہ	X
147	probably	شاید	X
148	produce	پیدا	X
149	provide	فراہم	X
150	punish	سزا	X
151	quiet	خاموش	X
152	rain	بارش	X
153	raise	بلند	X
154	ready	تیار	X
155	red	لال	X
156	remember	یاد	X
157	return	واپس	X
158	ride	سواری	X
159	rude	بدتمیز	X
160	salt	نمک	X
161	salty	نمکین	X
162	search	تلاش	X

163	sell	فروخت	X
164	sentence	سزا	X
165	shade	رنگ	X
166	shadow	سایه	X
167	sharp	تیز	X
168	sheet	چادر	X
169	shine	چمک	X
170	shop	دکان	X
171	side	کنارہ	X
172	similar	ملنے جاتے	X
173	since	بعد سے	X
174	slow	سست	X
175	small	چھوٹا	X
176	smile	مسکرایا	X
177	Smoking	تمباکو نوشی	X
178	So	تو	X
179	sound	آواز	X
180	speed	رفتار	X
181	spell	جادو	X
182	spend	خرچ	X
183	spring	موسم بہار	X
184	star	ستارہ	X
185	structure	ساخت	X
186	substance	مادہ	X
187	successful	کامیاب	X
188	sugar	چینی	X
189	summer	موسم گرما	X
190	swim	تیر	X
191	table	میز	X
192	thing	چیز	X
193	though	اگرچہ	X
194	toe	پنچہ	X
195	tooth	دانت	X
196	tree	درخت	X
197	twice	دو بار	X

198	try	کوشش کرنا	X
199	under	نیچے	X
200	vegetable	سبزی	X
201	voice	آواز	X
202	warm	گرم	X
203	way	راستہ	X
204	wedding	شادی	X
205	week	ہفتہ	X
206	welcome	خوش آمدید	X
207	while	جبکہ	X
208	white	سفید	X
209	wide	وسیع	X
210	wine	شراب	X
211	Winter	موسم سرما	X
212	Wire	تار	X
213	Wise	دانشمندانہ	X
214	Wish	خواہش	X
215	Worst	بدترین	X
216	Year	سال	X
217	Young	جوان	X

Source: Comparación del Léxico Básico del Español, el Inglés y el Urdu

Figure 4
Terms derived from Arabic languages.

	english meaning	term in urdu	arabic
1	abolish	ختم	X
2	act	عمل	X
3	accurate	صحیح	X
4	add	شامل کرنا	X
5	adequate	مناسب	X
6	after	بعد	X
7	age	عمر	X
8	agree	متفق ہونا	X

9	air	ہوا	X
10	all	تمام	X
11	angry	غصہ	X
12	answer	جواب	X
13	appear	ظاہر	X
14	approve	منظور	X
15	area	علاقہ	X
16	army	فوج	X
17	attack	حملہ	X
18	aunt	خالہ	X
19	away	دور	X
20	bath	غسل	X
21	beer	شراب	X
22	begin	شروع	X
23	besides	علاوہ	X
24	body	جسم	X
25	book	کتاب	X
26	border	سرحد	X
27	borrow	قرض	X
28	build	تعمیر	X
29	busy	مصروف	X
30	by	قریب	X
31	careful	احتیاط سے	X
32	careless	لا پرواہ	X
33	carry	اٹھانا	X
34	central	مرکزی	X
35	certain	بعض	X
36	chair	کرسی	X
37	chance	موقع	X
38	choice	انتخاب	X
39	choose	انتخاب کرنا	X

40	circle	دائرہ	X
41	class	جماعت	X
42	clean	صاف	X
43	clear	واضح	X
44	coin	سکہ	X
45	collect	جمع	X
46	common	عام	X
47	compare	موازنہ	X
48	complete	مکمل	X
49	condition	شرط	X
50	continue	جاری	X
51	corn	مکئی، کارن	X
52	contain	شامل	X
53	country	ملک	X
54	dance	رقص	X
55	dangerous	خطرناک	X
56	decide	فیصلہ	X
57	depend	انحصار	X
58	develop	ترقی	X
59	different	مختلف	X
60	difficult	مشکل	X
61	direction	سمت	X
62	dot	نقطہ	X
63	drop	قطرہ	X
64	duty	فرض	X
65	earn	کمانا	X
66	easy	آسان	X
67	education	تعلیم	X
68	effect	اثر	X
69	empty	خالی	X
70	end	آخر	X

71	enjoy	لطف اندوز	X
72	enter	داخل بونا	X
73	entrance	داخله	X
74	event	واقعه	X
75	exact	قطعی	X
76	exam	امتحان	X
77	example	مثال	X
78	expect	توقع	X
79	explain	وضاحت	X
80	extremely	انتہائی	X
81	face	چہرہ	X
82	fact	حقیقت	X
83	fake	جعلی	X
84	famous	مشہور	X
85	far	دور دراز	X
86	father	والد	X
87	fault	غلطی	X
88	fear	خوف	X
89	feel	محسوس	X
90	fever	بخار	X
91	finish	ختم	X
92	fit	موزوں	X
93	floor	فرش	X
94	force	مجبور	X
95	foreign	غیر ملکی	X
96	forgive	معاف	X
97	freeze	منجمد	X
98	further	مزید	X
99	future	مستقبل	X
100	general	عام	X
101	get	حاصل	X

102	gift	تحفہ	X
103	give	دینا	X
104	goodbye	الوداع	X
105	grave	قبر	X
106	great	عظیم	X
107	half	نصف	X
108	hard	مشکل، سخت	X
109	hate	نفرت	X
110	healthy	صحت مند	X
111	hear	سن	X
112	hello	آداب	X
113	help	مدد	X
114	helper	مددگار	X
115	high	اعلیٰ	X
116	hobby	شوق	X
117	house	مکان	X
118	hurry	جلدی	X
119	idea	خیال	X
120	important	اہم	X
121	increase	اضافہ	X
122	introduce	متعارف کرانا	X
123	invent	ایجاد	X
124	invite	مدعو کرنا	X
125	island	جزیرہ	X
126	join	شامل ہونا	X
127	just	صرف	X
128	key	اہم	X
129	kill	قتل	X
130	kind	قسم	X
131	know	معلوم	X
132	last	آخری	X

133	lately	حال بی میں	X
134	later	بعد	X
135	lead	قیادت	X
136	lend	قرض دینا	X
137	lesson	سبق	X
138	let	اجازت دینا	X
139	letter	خط	X
140	long	طویل	X
141	look	نظر	X
142	love	محبت	X
143	luck	قسمت	X
144	main	اہم	X
145	map	نقشہ	X
146	mean	مطلب	X
147	medicine	دوا	X
148	meet	ملاقات	X
149	member	رکن	X
150	mention	ذکر	X
151	method	طریقہ	X
152	mind		X
153	miss	امتحان	X
154	mistake	غلطی	X
155	modern	جدید	X
156	moment	لمحہ	X
157	more	مزید	X
158	morning	صبح	X
159	most	زیادہ	X
160	move	حرکت کرنا	X
161	must	ضروری	X
162	nation	قوم	X
163	nature	نوعیت	X

164	near	قريب	X
165	nearly	تقريباً	X
166	need	ضرورت	X
167	news	خبر	X
168	newspaper	اخبار	X
169	noble	عظيم	X
170	nose	ناک	X
171	notice	اطلاع	X
172	now	اب	X
173	number	تعداد	X
174	obey	اطاعت	X
175	object	اعتراض	X
176	ocean	بحر	X
177	office	دفتر	X
178	often	اکثر	X
179	only	صرف	X
180	opposite	برعکس	X
181	order	حکم	X
182	page	صفحہ	X
183	paper	کاغذ	X
184	parent	والدين	X
185	part	حصہ	X
186	pass	گزرنا	X
187	past	گزشتہ	X
188	pay	ادا	X
189	peace	امن	X
190	pen	قلم	X
191	perfect	کامل	X
192	person	شخص	X
193	picture	تصویر	X
194	pocket	جیب	X

195	point	نقطه	X
196	poor	غریب	X
197	popular	مقبول	X
198	position	مقام	X
199	possible	ممکن	X
200	power	طاقت	X
201	price	قیمت	X
202	private	ذاتی	X
203	prize	انعام	X
204	problem	مسئله	X
205	promise	وعدہ	X
206	proper	مناسب	X
207	protect	حفاظت	X
208	public	عوامی	X
209	pupil	طالب علم	X
210	queen	ملکہ	X
211	question	سوال	X
212	quick	فوری	X
213	quite	قطعی	X
214	real	اصلی	X
215	really	واقعی	X
216	receive	وصول	X
217	rent	کرایہ	X
218	repair	مرمت	X
219	reply	جواب	X
220	result	نتیجہ	X
221	rich	امیر	X
222	rule	حکمرانی	X
223	ruler	حکمران	X
224	rush	جلدی	X
225	safe	محفوظ	X

226	serve	خدمت	X
227	sex	جنس	X
228	shape	شکل	X
229	share	حصہ	X
230	sheer	صرف	X
231	shoulder	کندھے	X
232	sick	بیمار	X
233	silly	حماقت	X
234	silver	چاندی	X
235	sister	بہن	X
236	sit	بیٹھ	X
237	skill	مہارت	X
238	skin	جلد	X
239	slip	پہسلنا	X
240	smell	بو	X
241	smoke	دھواں	X
242	sock	موزا	X
243	soon	جلد ہی	X
244	sorry	معاف کرنا	X
245	south	جنوب	X
246	space	خلا	X
247	special	خصوصی	X
248	spring	موسم بہار	X
249	square	مربع	X
250	stamp	مہر	X
251	start	شروع	X
252	state	ریاست	X
253	step	قدم	X
254	store	نخیرہ	X
255	storm	طوفان	X
256	strange	عجیب	X

257	strong	مضبوط	X
258	study	مطالعه	X
259	subject	مضمون	X
260	sudden	غير متوقع	X
261	suitable	مناسب	X
262	summer	موسم گرما	X
263	support	حمایت	X
264	sure	یقین	X
265	surprise	حیرت	X
266	taste	ذائقه	X
267	terrible	خوفناک	X
268	test	امتحان	X
269	therefore	لهذا	X
270	threat	خطرہ	X
271	tidy	صفائی	X
272	title	عنوان	X
273	tool	آلہ	X
274	trouble	مصیبت	X
275	trust	اعتبار	X
276	turn	باری	X
277	type	قسم	X
278	use	استعمال	X
279	useful	مفید	X
280	usual	حسب معمول	X
281	usually	اکثر	X
282	visit	دورہ	X
283	wait	انتظار	X
284	waste	ضائع	X
285	weather	موسم	X
286	weight	وزن	X
287	weird	عجیب	X

288	west	مغرب	X
289	why	کیوں	X
290	wild	جنگلی	X
291	win	جیت	X
292	window	کھڑکی	X
293	wine	شراب	X
294	woman	عورت	X
295	wonder	تعجب	X
296	word	لفظ	X
297	world	دنیا	X
298	worry	فکر	X
299	zero	صفر	X

Source: Comparación del Léxico Básico del Español, el Inglés y el Urdu

Analysis

The basic vocabulary of Urdu language is the corpus of this study. This basic vocabulary was obtained through the translation of the basic vocabulary of English²⁵ into Urdu language. An etymological analysis of the corpus was performed which had revealed that the basic vocabulary of Urdu language is basically composed of Sanskrit, Prakrit, Persian, Arabic and English terms with a slight influence from Turkish language.²⁶ The English language influence is however, without any doubt, a more recent historical event. After performing the etymological analysis, the terms with etymology derived from Persian and Arabic were extracted with the purpose of obtaining the exact density of Persian and Arabic elements which compose the basic vocabulary of Urdu language. The etymological analysis of the basic vocabulary of Urdu yielded the following results:

Figure 5

Percentages of terms from Persian and Arabic present in Urdu language

	NUMBER OF TERMS	PERCENTAGE
Terms with etymology in Persian	217	21.7%
Terms with etymology in Arabic	299	29.9%
TOTAL IMPACT	526%	52.6%

The results reflect that 21.7% of the basic vocabulary of Urdu language has derived from Persian language, while 29.9% has derived from Arabic. We can then conclude that more than half of the basic vocabulary of Urdu language derives from Persian and Arabic languages.

Conclusion and Future Work

The lexicon is a register of terms that are accumulated in the speaker's brain. Its development from birth, evolution and renewal, new word formation, borrowing from other languages are subjects studied by the linguistics branch called word formation which analyzes the previously mentioned mechanisms. In this research, etymological sources have been applied through a diachronic approach by extracting the elements of Persian and Arabic present in the basic vocabulary of Urdu language, through its definition, and measure of these elements the exact lexical density was obtained utilizing a similar method as that of Fang and Jing.²⁷

A previous study revealed the lexical density of Sanskrit and Prakrit. In the present one, we have obtained the lexical density of Persian and Arabic. This *modus operandi* has a particular reason to be. The Sanskrit and Prakrit elements were studied together since the language (Urdu-Hindi) was unified at that time. Reforms were undertaken in Hindi language to rid it from Persian

and Arabic elements based on religious grounds since the texts of the Rig Veda were written in Sanskrit, while Sanskrit elements were being extensively borrowed. The opposite phenomenon was progressively occurring in Urdu language; a large percentage of Sanskrit elements were removed, as well as those rustic or obsolete elements of other origins. This resulted in more than half of the modern basic lexicon of Urdu being borrowed from Persian and Arabic which we have studied in this research, creating the split of both languages Hindi and Urdu.

Since the original study revealed substantial elements of Sanskrit, Prakrit, Persian, Arabic and English and the first four elements have already been studied, a future research will be concerned with those elements of English language which make up the composition of the basic vocabulary of Urdu language. The density of the English terms will be the object of a future study.

Notes and References

- 1 Tariq Rahman. Urdu and the Muslim Identity: Standardization of Urdu in the Eighteenth and Nineteenth Centuries. *Annual of Urdu Studies*. Vol.25. (2008) 24. See also R. D. King. The poisonous potency of script: Hindi and Urdu. *International journal of the sociology of language* (2001). 43-60. George Cardona. Indo-Aryan languages. In *The World's Major Languages*, Bernard Comrie (ed.), (New York: Oxford University Press. 1987a) 440-447. S. K. Chatterji. Indo-Aryan and Hindi. (Calcutta: K.L. 1969). Mukhopadhyay. 167. Kerrin Dittmer. Die indischen Muslims und die Hindi-Urdu Kontroverse in den United Provinces. (Wiesbaden: Otto Harrassowitz. 1972) 263-272. Yamuna Kachru. Hindi-Urdu. In *The World's Major Languages*, Bernard Comrie (ed.), (New York: Oxford University Press 1987) 470-489. Ashok Kelkar. *Studies in Hindi-Urdu*. Poona: Deccan College Postgraduate and Research Institute. (1968) 49-63. Christopher R. King, *One Language, Two Scripts: the Hindi Movement in Nineteenth Century North India*. (Bombay: Oxford University Press, 1994) 4.

-
- 2 Christopher R. King. *One Language, Two Scripts: the Hindi Movement in Nineteenth Century North India*. (Bombay: Oxford University Press, 1994). Also see King, Christopher R. *One Language, Two Scripts: the Hindi Movement in Nineteenth Century North India*. (Bombay: Oxford University Press, 1994).
 - 3 Maria Isabel.Maldonado Garcia. *The Urdu language Reforms*. *Almas Urdu Research Journal* (2013) Vol.:14.
 - 4 *The Mughal daynasty in India started from Zaheer-ud-Din Babur in 1526 and collapsed with the British occupation in 1857.*
 - 5 Rahman, Tariq. *Language, Religion and Politics. Urdu in Pakistan and North India*. *Revue des mondes musulmans et de la Méditerranée*, 124, (2008). 90.
 - 6 King,R. D. *The poisonous potency of script: Hindi and Urdu*. *International journal of the sociology of language* (2001). 43-60.
 - 7 Ibid
 - 8 Ibid
 - 9 Schimmel, Anna Marie. *Classical Urdu literature from the beginning to Iqbal*. *Otto Harrassowitz Verlag*. (1975) 126.
 - 10 *Encyclopaedia Britannica* (online version). <http://www.britannica.com/EBchecked/topic/384749/Mir-Ali-of-Tabriz> Accessed 5th September, 2014.
 - 11 Maria Isabel Maldonado Garcia & Mustapha Yapici. *Common Vocabulary in Urdu and Turkish Languages. A Case of Historical Onomasiology –Journal of Pakistan Vision* Vol.15. No. 1. (2014) 3.
 - 12 Christopher R. King, *One Language, Two Scripts: the Hindi Movement in Nineteenth Century North India*. (Bombay: Oxford University Press, 1994) 4. See also King, R. D. *The poisonous potency of script: Hindi and Urdu*. *International journal of the sociology of language*, (2001). 45.

-
- 13 Anna Marie Schimmel. *Classical Urdu literature from the beginning to Iqbāl*. Otto Harrassowitz Verlag. (1975) 127.
 - 14 Figure taken from Maldonado García, María Isabel. *Comparación del Léxico Básico del Español, el Inglés y el Urdu*. (Unpublished doctoral dissertation). (UNED, Madrid. 2013).
 - 15 Skjærvø, Prods Oktor. *An Introduction to Old Persian*. (2002): 13.
 - 16 Pourjafar, M. Reza, and Ali A. Taghvae. "Indo-Iranian Socio-Cultural Relations at Past, Present and Future (with Special Reference to Architecture of Mughals' Gurkanids' Period) (2006): 78.
 - 17 Tariq Rahman,. *Urdu and the Muslim Identity: Standardization of Urdu in the Eighteenth and Nineteenth Centuries*. *Annual of Urdu Studies*. Vol.25. (2008) 99.
 - 18 Nehru, J. *An Autobiography*, London: The Bodley Head, 1989, p 149.
 - 19 *Ibid*, 151.
 - 20 Figure created with information from Ethnologue and Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (West Semitic is not represented on this table). Available online at <http://wals.info>, Accessed on 2014-03-24.
 - 21 R.D.King. (2001) 46.
 - 22 Tariq Rahman. *Urdu and the Muslim Identity: Standardization of Urdu in the Eighteenth and Nineteenth Centuries*. *Annual of Urdu Studies*. Vol.25.(2008) 90.
 - 23 Tariq Rahman states (in *Urdu and the Muslim Identity: Standardization of Urdu in the Eighteenth and Nineteenth Centuries* (2008) 83.) that Persian was discontinued in India as an administrative language in 1834. However, Christopher R. King (in *One Language, Two Scripts*, 9) states that it occurred in 1837.

-
- 24 URDU DICTIONARY (2011): Urdu Encyclopedia. Islamabad, Ministry of Science and Technology. Accessed originally on 13th September, 2013 and for the purpose of this paper on 28th August, 2014. <http://www.urduencyclopedia.org/urdudictionary>
 - 25 Basic vocabulary of English language, 1000 words http://simple.wikipedia.org/wiki/Wikipedia:List_of_1000_basic_words.
 - 26 María Isabel Maldonado García. Comparación del Léxico Básico del Español, el Inglés y el Urdu. (Unpublished doctoral dissertation-UNED, Madrid. 2013) 500.
 - 27 Alex Chengyu Fang & Cao Jing. A Corpus-Based Quantitative Survey of the Etymological Composition of Contemporary British English: The Case of Latin, Greek and French. *Glottology* 3, no. 1 (2010): 53-68.