Lexical affinities between Urdu and French

Ahmad Ali Khan
Institute of Languages, University of the Punjab
Lahore, Pakistan
Catullus2001pk@yahoo.com

Abstract

This paper provides a comparative lexical survey of Urdu and French languages in a historical context. Especially, it presents a detailed etymological analysis of selected Urdu and French words to demonstrate their common ancestral origin. Despite the fact that Urdu and French are genetically related languages, their comparative study is one of the more poorly studied areas till now. This study also provides a historical background of both languages stressing the parallel development in their linguistic evolution.

Keywords: Comparative method, language families, genetic relatedness, cognates, proto-languages, etymology.

Resumen

Este artículo revisa el léxico comparativo de las lenguas urdu y francés en un contexto histórico. Especialmente, se presenta un análisis detallado de Urdu etimológico seleccionado y palabras francesas para demostrar su origen ancestral común. A pesar del hecho de que el urdu y el francés son lenguas emparentadas genéticamente, su estudio comparativo es una de las zonas más poco estudiadas hasta ahora. Este estudio también proporciona un fondo histórico de los dos idiomas hincapié en el desarrollo paralelo en su evolución lingüística.

Palabras clave: método comparativo, familias de lenguas, relación genética, cognados, proto-lenguas, etimología.

Introduction

All human languages resemble each other for the following three reasons:

1. They share some common features because these features show universal and/or typological characteristics.

2. They resemble each other because they are in contact with each other and thus influence each other.

3. Languages share common features because they have inherited common linguistic features from a proto- language, which means that the two languages are genetically related. In this article, we will be dealing with this third reason.

For more than two hundred years genetic linguistics has been one of the most productive and appealing research area among linguists (Gell-Mann, Peiros & Starostin, 2009). It is true that many great advances have been made in this field, but due to lack of proper coordination among experts and widely varying degrees of our knowledge about different linguistic areas in the world, a lot still needs to be done. That is why we do not see much comparative work on languages which are rather remotely related, e.g. Urdu and German, Hindi and Dutch, French and Punjabi etc.

However, in the last few decades we have seen a fundamental change in the techniques of comparative linguistics. A long period of research in comparative studies with a wide range of language families has laid the foundation for the study of genetic relationships between remotely related languages or language groups.

Language relatedness: the comparative method

Practically the method used for proving genetic relationship for remotely related languages is the same as used for any language family relationship, near or not. It is the basic tool used to establish genetic relationship and it is a set of procedures and techniques developed over more than a century (Joseph & Janda, 2003). This method allows us to recover linguistic data (which may be morphological, phonological, lexical, syntactic, etc.) from usually unrecorded and earlier stages of a group of related languages, or they may be conditions, rules and constraints of grammar. The method proceeds by comparing the cognate vocabulary from

two or more related languages from a language family. Thus by comparing systematically we get sets of regularly corresponding forms from which we can derive an antecedent form and determine its place in the proto-language. This (Joseph & Janda, 2003) always involves by beginning with basic cognate vocabulary, establishing similar sound correspondences and reconstructing proto-syntactic and phonological system. The comparative method proceeds in various stages, thus a complete comparative treatment of a linguistic family would include reconstruction of phonology, syntax, semantics and morphology (Joseph & Janda, 2003).

It is assumed that proof of linguistic affinity relies mostly on examination of analogies and similarities among the languages compared, but in fact this is not enough, because the similarities can be due to borrowing, chance, language universals, onomatopoeia etc, as well as due to common linguistic inheritance. For this reason, anyone who is interested to demonstrate a genetic relationship among languages has to prove that the evidence given for a hypothesized relationship cannot be easily explained by factors other than genetic. Considering the problems which certain claims can cause about the proposed genetic relationship regarding the remotely related languages, it is important to apply the methodological procedures and principles (lexical comparison, basic vocabulary, glottochronology, sound correspondences, grammatical evidence, borrowing etc) carefully. Thus, it is necessary to rely on regularly patterned grammatical structures combined with sound correspondences in cognate vocabulary and at the same time eliminate other possible similarities such as chance, borrowing, typological universals, onomatopoeia etc (Campbell & Poser, 2008).

Historical linguistics is the historical study of language change and development. Its results are directly relevant to comparative linguistics, because only by taking into account the history of languages can we understand why some of them share some of the features they do. Therefore, in order to show the genetic relatedness of Urdu and French languages, we give a short history of the Indo-European language family.

The term Indo-European was first introduced in 1816 by the German linguist Franz Bopp to designate a set of languages in Europe and Asia (including northern India with Iran, Afghanistan, Pakistan and Bangladesh) which showed notable linguistic affinities. Greek, Latin, Sanskrit, Irish, Gothic, Bulgarian, Hittite, etc. show striking correspondences in their vocabulary, syntax and phonology — which link them together. The source language, generally called Proto-Indo-European, was spoken some 6,500 years ago and has given rise to several hundred languages, in ten major branches.

Currently the Indo-European family includes Indo-Iranian languages including Sanskrit, Hindi/Urdu and Persian, Greek, Baltic languages such as Lithuanian and Latvian, Celtic languages such as Breton, Welsh or Scottish Gaelic, Latin languages or say Romance languages such as Spanish, Italian, French, Catalan, Portuguese, Germanic languages such as English, German, Dutch and Scandinavian languages, Slavic languages such as Russian, Polish and Serbian. The Indo-European languages consisting of around 450 languages and dialects out of which about half belong to the Indo-Aryan subfamily are spoken by almost 3 billion people. (Ethnologue, 2015) This family is significant to the study of historical linguistics as it has some of the longest written records dating since the Bronze Age in the form of Mycenaean Greek, Sanskrit and Anatolian Languages (Clackson, 2007).

According to Lockwood (1972) the grammar of the Indo-European languages has developed on the following principles: a distinction between animate and inanimate objects being further divided into masculine, feminine and neuter; nouns and adjectives were declined as seen in Sanskrit, Greek and Latin languages and there were eight cases viz., nominative, vocative, accusative, instrumental, genitive, dative, ablative and locative. Each case is identified by a special suffix. The declension of adjectives was in principle identical with that of nouns. Adjectives agreed with nouns in number and gender. The distinction of two numbers i.e. singular and plural is also evident in the grammar. The verbal system has three basic tenses, namely, present, past and future and these can be categorized either as transitive or intransitive verbs. The past tense transitive verb agrees with the object in gender and number.

Evidence uniting the Indo-European languages

As mentioned above, the basis of an Indo-European family and for linking the various languages listed above is a set of linguistic correspondences at various levels of morphology, phonology, syntax and lexicon. An important feature of these correspondences is that those involving sounds are internally consistent and regular, the cognate lexicon and morphemes also exhibit correspondence in meaning and grammatical functions across the various languages. The matching of form and meaning is really significant, especially when it involves recurring irregular patterns; for example, the substitution evident in the English adjective comparison pair good/better is matched by German gut/besser; and similarly khub/behter in Persian. Likewise the verb 'be' is irregular in almost all Indo-European languages. Examples like these can be

multiplied to give conclusive evidence of a genetic relationship among the various languages of the Indo-European family.

Historical background of Urdu and French languages

Aitchison (2001) has outlined three theoretical possibilities of language change. The first possibility is slow decay, where a standard language is used in a formal setting but the same variety changes its form in an informal setting. Aitchison (2001) states that many scholars were of the opinion that European languages were on the decline because they were gradually losing their word endings. In this regard, she cites the popular German writer Max Muller who asserted that "The history of all Aryan languages is nothing but a gradual decay" (Aitchison 2001). The second possibility is that languages may be slowly evolving to a more efficient state by becoming streamlined and sophisticated. We may be witnessing the survival of the fittest with existing languages adapting to the needs of the times. Thirdly, a language remains in a substantially similar state from the point of view of progress or decay dominated by certain general laws.

Now if we consider the history of Urdu and French languages, we see that both Aitchison and Max Muller are right in their theses about language change, since both Urdu and French were initially the vernacular and substandard version of standard languages before evolving 'to a more efficient state.'

Urdu

Urdu is the cultural, religious and literary language of Muslims in Pakistan, India, Bangladesh, UAE and other parts of the world including the, the United Kingdom, United States, Germany, France and Spain (Brown & Ogilvie, 2009). Urdu is not only the national language of Pakistan, but also the state official language of Jammu and Kashmir.[i] It is also one of the Schedule VIII languages of India,[ii] and the second official language of UP, Andhra Pradesh and Bihar in India (Brown & Ogilvie, 2009). According to Ethnologue (2015) there are between 60 and 70 million native speakers of Urdu: there were 52 million in India per the 2001 census (The Census of India, 2001) and 10 million in Pakistan (The Census of Pakistan, 1998). At the colloquial level Urdu and Hindi are almost identical in their vocabulary as well as in grammar; but they differ at an advanced or technical level where Urdu gets more Arabicized or Persianized and Hindi gets Sanskritized (Comrie, 2009). If Hindi and Urdu are considered as the

same language, their number of speakers becomes the fourth largest in the world, coming after Chinese, English and Spanish (Ethnologue, 2015).

Urdu is an Indo-Aryan language which belongs to the Indo-Iranian branch of the Indo-European family of languages (Ethnologue, 2015). Sanskrit is considered as the mother of all modern Indo-Aryan languages (Desai & Ramsay-Brijball, 2004). In time, Sanskrit gave birth to many dialects which were known as the Prakrits (a term used for spoken dialects) and then the Apabhramshas, the dialects considered as corruption of the norm. It is from the latter that the Modern Indo -Aryan languages such as Urdu, Punjabi, Gujarati, Hindi and Bengali, etc. derived (Comrie, 2009). The term Urdu derives from the Turkish word Ordu meaning camp or army.

Historically, Urdu developed in a language contact situation over a long period from 1100 A.D. or earlier. Specimens of Hindvi in the early formative period are found scattered in the Nath Panthi literature, early Sufis of North India, Amir Khusro, Nanak, Kabir, Baba Farid, and other poets (Brown & Ogilvie, 2009). Amir Khusro (1236–1324) shows a distinct earlier form of Urdu, or Hindwi as he calls it. However, according to Khan (1969) there is no evidence that the language was in continuous use from 1200 to1650 except Bikat Kahani by Afzal, which appeared 300 years after Amir Khusro's writings. It is therefore not possible to reconstruct a continuous history of the development of Urdu.

Urdu is regarded as a pluri-centric language that shows different linguistic features. It is generally recognized that Urdu and Hindi share a common grammatical system. They differ mainly in their writing systems, in their lexicon borrowed from Sanskrit or Persian and Arabic resources, and the minor aspects of syntax (Brown & Ogilvie, 2009).

French

French is a Romance language derived from Latin which belongs to the Indo-European Family of Languages. It is spoken today on all continents by about 274 million people including 76 million to 77 million native speakers.[iii] It is spoken in France and in Belgium, Canada, Luxembourg, Switzerland and 51 other countries, mainly located in Africa, which were once part of the former French colonial empire. It is one of the six official languages and two working languages (with English) of the United Nations, and official language or working language of many international and regional organizations, including the European Union. During the later Middle Ages and Early Modern era French remained a language of European high gentry and

international diplomacy (Walter, 1994). It is still the second-most widespread international language after English and the language most taught after English (La Langue Française dans le Monde, 2014). French is spoken in 51 countries (Ethnologue, 2015).

Resulting from the evolution of Latin to the Gallo-Roman in the first millennium of the Christian era, French became a legal and administrative language with the Ordinance of Villers-Cotterets in 1539 (Fagyal, Kibbee & Jenkins, 2006). French is a Romance language which means that its grammar and most of the vocabulary comes from Latin; but its morpho-syntax and especially its phonology have undergone such radical changes since Latin times that it looks very different from its sister languages i.e. Spanish, Italian and Portuguese. In the Middle Ages, the French language is made of a multitude of dialects that vary considerably from one region to another. These were the dialects of Oïl in the North and Oc (Occitan) dialects in the South, both Oc and Oïl meaning "yes" (Comrie, 2009). With the establishment and consolidation of the Capetian monarchy, it is the dialect of Oïl which gradually got ascendency. Thus in the Middle Ages, France had no linguistic unity, since the beginning of the IX the century the Latin language was only a "sacred" or intellectuals' language and France was fragmented by different dialects.

It is sometimes referred to as the Francien, this proto-French dialect as langue d'oïl spoken in Ile-de-France in the Middle Ages, is the origin of today's French. The French spoken today takes its name from the ancient Franceis, which evolved slowly to a supra-regional dialect from the eleventh and twelfth centuries from Franceis to François to French.

The other langue D'oïl disappeared little by little as French took the prestige but also because Langue D'oïl was very close to François. The François is enriched quickly thanks to the contribution of other dialects and as Paris became important, more people across the country flocked there, bringing with them their language variant. Paris and its region are the historic cradle of this Franceis which soon was enriched by the contribution of Norman, Picard, Burgundian and other dialects of neighboring dialect of Oïl.

The Romance version of the Oaths of Strasbourg of 842 is the first text written in the langue D'oil, derived from the Lower Latin and remodeled as a result of the establishment of Germans, mainly Franks (hence the name of the French) in northern Gaul.

The period between the end of the eleventh century to the beginning of fourteenth century is a glorious period of the medieval French. French became an international language, spoken in all over Europe, thus becoming the "lingua franca "of the world. It became the language of international law and diplomacy.

Norman French was introduced in England by Norman conquerors in 1066. The Norman occupation of England lasted more than three hundred years. Therefore, English vocabulary has strong influence from French: A substantial proportion of the English lexical content comes from the French Language. It had a strong influence on the medieval Italian literature as many Italian writers wrote in French: Martino da Canale wrote his chronicle of the Venetians; the famous Florentine philosopher Brunetto Latini (1220-1294) wrote in French language (langue d'oil) the Book of Treasure; Marco Polo dictated his travel stories in French in the prison of Genoa and The famous German philosopher and scholar Gottfried Wilhelm Leibniz (1646-1716), wrote mostly in French (Walter, 1994).

Child of Latin, French replaces Latin as an international language in seventeenth century before leaving its place to English since the end of World War II. The main cause of this French hegemony lies in the power of the French state at the time. In 1685, Pierre Bayle (1647–1706) wrote that French is "the communication point of all the peoples of Europe." Since the eleventh century French Literature has set standards and provided models not only for the Western world but whose influence spread in the nineteenth and twentieth centuries beyond the confines of Europe to Africa.

Theoretical Background of Urdu-French Relatedness

Filippo Sassetti, a merchant from Florence, was on a business trip to Goa for about five years from 1583. He was struck by the similarities between Italian (or its parent language Latin) and Sanskrit of India. Not being a scholar, he did not try to find the reasons. But his preliminary observations posed an important question before linguists – why there should be similarities between different languages? (Marcucci, 1855).

The discovery of Sanskrit by European scholars towards the close of the 18th century was the starting point of the scientific study of comparative linguistics. It was observed that in both vocabulary and grammar Sanskrit was remarkably similar to the majority of the languages in Europe and particularly in grammar, to the classical languages. The only theory that could

explain these fundamental similarities was that all the languages in question were derived from a common parent language (Emeneau, 2000).

The relationship was further explored by William Jones (1786) who supported the idea that Latin, Sanskrit, Greek, and several other Indo-European languages had "sprung from some common source (Desai & Ramsay-Brijball, 2004). In his historic lecture before the Asiatic Society in Calcutta in 1786, he also emphasized that the similarities between the classical languages of Europe and Sanskrit had not only similar shapes of words, but also similar grammatical structures. In his *Comparative Grammar* (published between 1833 and 1852) the German linguist Franz Bopp expounded the similarities between verbal systems of Greek, Sanskrit and Latin with other Indo-European language to show their genetic relationship. Further studies were made by Jakob Grimm (1822) who worked on sound correspondences between Germanic and other Indo-European languages and by August Schleicher (1861) who introduced the classic genealogical tree-diagrams to show the genetic relatedness of languages (Mukherjee, 1997-1998).

As stated earlier that languages have been traditionally classified in terms of the genetic relationships that they exhibit. It has been proved useful to group languages into language families. Within a given family, languages can be "parents" of other languages, "siblings" of one another, and so forth. These relationships are expressed through successive branching of a family tree reflected in figure one.

Fig. 1. Proto-Indo-European Languages
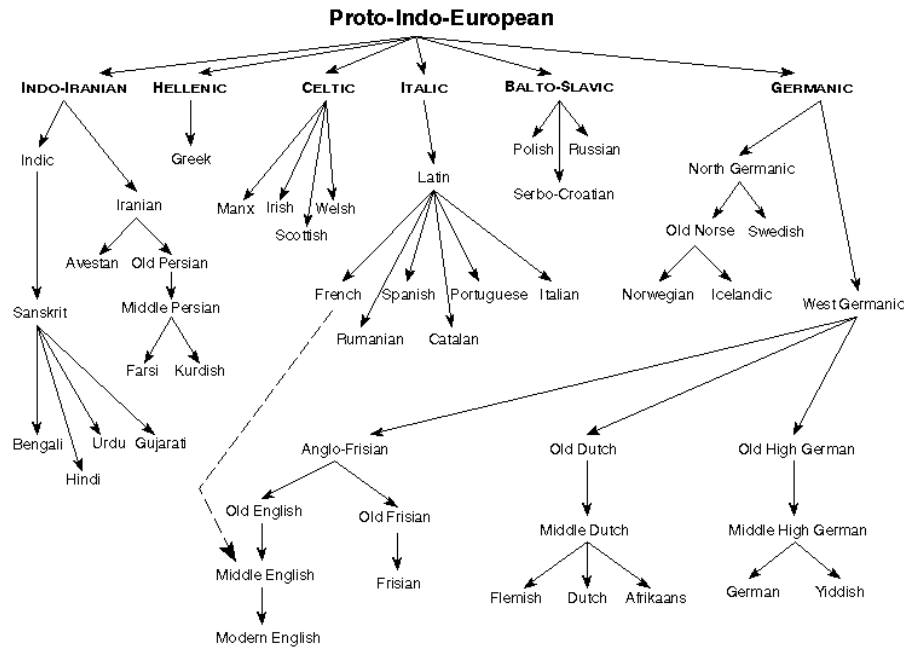
Source: Jack Lynch, Rutgers [iv]

Urdu (through Sanskrit) and French (through Latin) are related at the Indo-European level; that is, if the Indic languages were thought as being "sister" languages, Italian would be a "cousin" language to them. The family-tree model of languages shows how languages change over time. Indo-European, for example, is said to have existed approximately 4000 BC or even earlier; Indo-Aryan languages to 1500-3000 BC and Italic to 1000 BC. To develop a family tree and establish genetic relationship among languages, various kinds of linguistic, historical and archeological evidence are examined.

There are basically three kinds of evidence to establish the relationship between the members of Indo-European or any other language family- grammatical similarities, cognate vocabulary, historical and archeological information.

Up to now it has been shown typologically and historically that French and Urdu have both sprung from the Indo-European family. Now their genetic relationship in linguistic terms will be explored by providing the required evidence from cognate vocabulary and grammatical similarities.

Lexical similarities between Urdu and French

Cognate Vocabulary

Cognates are defined as words in two languages that have a common etymology and therefore are similar more or less in meaning and form, however, sometimes the meaning changes in one language or another. The comparison of cognate vocabulary is the distinctive mark of the comparative method. Cognates are words that are passed down the family tree as languages change and develop and have proven extremely important for determining not just which languages are siblings within a language family but what the parent language of the sibling languages might have looked like.

The comparative method works best when vocabulary representing common human experiences is compared. Watkins (2000) lists many semantic categories containing words that were instrumental in developing the Indo-European family (examples from Modern English are used for purposes of illustration): for instance, verbs of existence (e.g. English *be*); qualitative adjectives (*old*, *new*, *thin*); numerals (*one*, *two*, *three*, etc.); pronouns (*I*, *me*, *you*, etc.); seasons (*winter*, *spring*, *summer*, *autumn*); body parts (*hands*, *nose*, *feet*, etc.); and so forth. A collection of 100 Swadesh (1971) words List (A list of basic vocabulary developed by the American linguist Morris Swadesh for comparative or historical linguistics) or 200 basic words is often used initially in cognate searches, the idea being that basic concepts are the least likely to have been borrowed (Swadesh, 1952). The advantage of comparing vocabulary such as this is that one can be assured that it will occur in almost any language. Culturally specific vocabulary has limited occurrence, and thus not suitable to the comparative method.

Urdu-French Cognates

The main source of most French words is Latin or rather Vulgar Latin, but this is not the case with Urdu. Sanskrit or its vulgar form Prakrit is not the main source of most Urdu words. Like English, Urdu has borrowed heavily from languages (Arabic and Persian) other than its parent language. However, both languages are similar in making phonetic alterations before borrowing words from the parent languages. Urdu, for instance, changed Sanskrit "sh" into "s" (desh→ des: country), "v" becoming "b" (vars→ baras: year), and "y" becoming "j" (yantra→ jantar: device). Sometimes short vowels were introduced to break up consonant clusters as in vars→ baras: year. Likewise in French many changes were made both in vowel and consonant sounds:

a. Many stressed vowels were diphthongized, so that Latin *pedem* (foot) became *pied, mel* (honey) *>miel, pira* (pear)*>poire,* and so on;

b. The voicing of k, p and t to g, b and d, between vowels, and the disappearance of t and c in this position, while p>v; so *securum>sûr ,vita>vie, ripa>rive;*

c. The 's' after a vowel and before a consonant disappeared which , lengthened the preceding vowel, and which is marked by a circumflex accent, as *in festa* (festival)*>fête, castellum* (castle)*>château;*

d.   The first element in many consonant clusters were either reduced or assimilated, so that *debita* (debt)*>dette and insula* (island*)>ile.* (Boyd-Bowman, 1980)

Following is a list of some selected Urdu and French cognates. The words in this list are more or less similar in meaning and form and they can be included in any list of most frequently used words.   Note that only native words are included i.e. words borrowed from another language are not included. For example Urdu has borrowed many words from Portuguese: انناّس (pineapple),   چابی(key), گرجا(church), کمرہ(room) etc and from English: بوتل(bottle), درجن (dozen), جنوری(January), اسکول(school), پستول(pistol) etc; now these words and many others like these are common words in Urdu and French, but they are not included for the above-mentioned reason. Further etymologies of some words are given for illustration to prove their genetic relatedness.

Methodology

The sample of our research consists of 50 pairs of shared words in both Urdu and French. These words are selected due to their apparent phonetic and semantic similarity. This similarity will be analyzed with reference to different linguistic aspects:

(i)   Etymology: Etymology of each shared lexical pair will be extracted down to the Indo-European lineage to prove the common ancestry of Urdu and French languages.

(ii)   Semantics: Meanings of these shared words will demonstrate whether the cognates are synonymic or not.

(iii)   Phonetics: Phonetic transcription of each pair will show the degree of phonetic similarity.

Data Presentation

The data is presented in the form of tables. First, the pairs of Urdu and French words are given in tables along with their semantic and phonetic similarity. Then etymological analyses are made, thus goes our etymological route: Urdu > Sanskrit (Prakrit) > Indo-European; Italian > Latin > Indo-European.

Table 1. French- Urdu Cognates

Sr. #	English	French	Urdu	French Phonetic Transcription	Urdu Phonetic Transcription
1	on foot	a pied	پیادہ	a ˈpjɛd	pjɛdah
2	Drink	boire	پینا	bwar	piːna
3	Boil	bouillir	ابالنا	bujir	ubalna
4	Hundred	cent	سو	sã	So
5	Thing	chose	چیز	ʃoz	ˈtʃiːz
6	Cry	cri	گریہ	kʁi	gʌrja
7	Weak	débil	دبلا	debil	dubla
8	Teeth	dent	دانت	dã	dãt
9	Two	deux	دو	dø	Do
10	God	dieu	دیوتا	djø	devta
11	Ten	dix	دس	dis	dʌs
12	Give	donner	دینا	dɔne	dena
13	Entrails	entrailles	انتّریاں	ãtʁaj	əntʌrjã
14	knee	genou	زانو	ʒənu	zanu
15	People	gens	جنتا	ʒã	dʒənta
16	drop of water	goutte	گھونٹ	gut	ghunt
17	Group	groupe	گروہ	grup	gəroh
18	Eight	huit	آٹھ	ˈɥi(t)	ath
19	Young	jeune	جوان	ʒœn	dʒvan
20	in french 'me' and in urdu 'i'	me	میں	mə	mɛ̃
21	Tit	mamelle	مما	mamɛl	mʌmma
22	Mother	mere	مادر / ماں	mamã	mã
23	My	mon	میرا	mɔ̃	mera
24	Dead	mort	مردہ	mɔʁ	murda
25	Die	mourir	مرنا	muʁiʁ	mʌrna
26	Middle	moyen	درمیان	mwajɛ̃	dʌrmijan
27	Nose	nez	ناس/ ناک	nez	nas/nak
28	Ship	navire	ناؤ	naviʁ	naʊ

29	Nine	neuf	نو	nœf	No
30	no, not	non/ne	نہ/ نہیں	nɔ̃	nahin/ na
31	name	nom	نام	nɔ̃	nɑm
32	New	nouveau	نیا	nuvo	nəja
33	nail in french; finger in urdu	ongle	انگلی /انگل	ɔ̃gl	ungli
34	frying pan	poêle	پتیلا	pwal	pətila
35	Foot	pied	پاؤں	pje	Paon
36	hair of body in french; hair of both head and body in urdu	poil	بال	pwal	bɑl
37	that/which	que	کہ	kə	Ke
38	what	que	کیا	kə	Kja
39	father	papa	باپ	papa	bɑp
40	king/prince	roi	راجہ	rwa	rɑdʒa
41	rustic	rustique	ئیروستا	rystik	rɔstai
42	chest	sein	سینہ	sɛ̃	Sina
43	seven	sept	تسا	sɛt	sɑt
44	snake	serpent	سانپ	sɛʀpɑ̃	sɑnp
45	(so) much	tant	اتنا	tɑ̃	itna
46	Three	trios	تین	trwa	tiːn
47	You	tu	تو	ty	Tu
48	widow	veuve	بیوہ	vœv	bevah
49	Live	vivre	جینا	vivr	dʒina
50	Voice	voix	آواز	vwa	ɑvɑz

Word Etymologies

After the comparison of the words, the etymologies were extracted.[v]

Table 2. Word etymologies

French	Urdu	Etymologies
Boire	پینا	The French word is derived from the Latin word "bibere" and the Urdu one is derived from the Sanskrit "pibati". The root etymology is the Indo-European *poi.
Cent	سو	Centum is the source of the French *cent* while Sanskrit *sata* is the source of the Urdu word.
Débil	دبلا	In Urdu this is from Prakrit which in turn derives from Sanskrit "durbala*."* In French the word comes from Latin "debilis" and the Indo-European.
Dent	**دانت**	PIE *h3d-nt- 'tooth'. In Urdu most probably from Prakrit through Sanskrit "dant" or perhaps from modern Persian "dandan."The Latin "dens, -tis" is the source of the French

		word.
Deux	دو	Proto-Indo-European word is "*du-o-h" which is the source of the Latin *duo* and Sanskrit *dvau / duvau* which in turn are the source of both French and Urdu respectively.
Dieu	دیوتا	PIE *dei-u-o- 'god. The Sanskrit word "deva" is the source for Urdu word and Latin "deus" is the source for French.
Donner	دینا	The French word is from the Latin word "donare" and the Urdu one is from the Sanskrit "da" which in turn from the Indo-European root *doh*.
Genou	زانو	The Urdu word is derived from the Persian "zanu" which in turn is derived from the Avestan *znu*. The French word is from the Latin "genu." The source of all these is the Indo-European *genu.
Huit	آٹھ	: PIE (Proto Indo-European) *h3(e)ktehj'eight'. In Urdu the word is from Sanskrit "asta, astdu", while in French from Latin "octo".
Jeune	جوان	PIE *h2iu-hien- 'who possesses vital force' > 'young. The French word is from Latin "iuvenis 'young man" and the Urdu one is from Persian which is ultimately from Avestan "yuuansm [acc.sg.], yunqm [gen.pl.]".
Mamelle	مما	Both Latin and Sanskrit words are "mamma and mama" respectively, while the Proto-Indo-European root is *mama.
Me	میں	The French word is from the Latin "me" and the Urdu one is from the Sanskrit *mam* and the original Indo-European root is *hme.
Mère	ماں/ مادر	PIE *meh2-tr- 'mother'. Sanskrit "mātár" and Latin "mater." The French word has changed from "medre" to mère.
Mourir	مرنا	PIE *mr-ie/o- 'to die', *mr-to- 'dead', *mr-uo- 'dead', *mr-ti- 'death'. The Urdu word has Sanskrit origin pr. mriyate (< *mr-ie~), aor. mar-Imf-, pf. mamara, caus. mardyati 'to kill', martave [inf.] 'to die', mrtd- 'died, dead'.
Moyen	درمیان	The Latin word "medius" is the source of the French one and the Sanskrit word "*mddhya*" is the source of the Urdu word. Both Sanskrit and Latin words are from the Proto-Indo-European *medhi-o. Note that the Urdu word میدان (play ground) is also from the same root.
Navire	ناؤ	PIE *neh-u- 'ship'. In Urdu the word came from Prakrit, but the actual source is the Sanskrit word "nau". In French the source is

		the Latin word "Navis".
Neuf/Nou veau	نیا	The French word is derived from the Latin word "novus" and the Urdu one is derived from the Sanskrit "nava". The root etymology is the Indo-European *neuo.
Nez	ناس/ ناک	The French word is derived from the Latin word "nasus" and the Urdu one is derived from the Sanskrit "nas." The original source is the Indo-European *Hnehs.
Nom	نام	PIE *h3neh3-mn, name. The derivation in French is from Latin "Nomen" and in Urdu from Persian which in turn is from Avestan "Nāman.".
Non/ne	نہ /نہیں	The Latin word "ne-" is the source of the French one and the Sanskrit word "na" is the source of the Urdu word. Both Sanskrit and Latin words are from the Proto-Indo-European *ne.
Ongle	انگلی /انگل	The Latin word "unguis" is the source of the French one and the Sanskrit word "anguli" is the source of the Urdu word. Both Sanskrit and Latin words are from the Proto-Indo-European *(o) nogh-.
Pied	پاؤں	PIE nom.sg. *pod-s. French is from Latin "pes, pedis' and Urdu from Sanskrit "pada".
Roi	راجہ	PIE *h3reg-e/o- 'to stretch, direct. The Sanskrit word "Raj 'to stretch, direct" is the source of the Urdu word, while the Latin word "Rego, -ere 'to direct, guide, govern" has given the French word.
Serpent	سانپ	PIE *serp-e/o- 'to creep'. In Urdu this is from Prakrit whose source is Sanskrit "*sarpati* 'to creep, crawl', *sarpa-* snake". In French the word comes from Latin "*serpens* snake, which from serpo, -ere 'to crawl'".
Tu	تو	PIE. *tu. In Urdu the derivation is from Sanskrit "továm" or from Persian "tuvam" and in French from Latin "tū."
Trois	تین	The Latin word "tres" is the source of the French one and the Sanskrit word "trayas" is the source of the Urdu word. Both Sanskrit and Latin words are from the Proto-Indo-European *tri.
Veuve	بیوہ	The Urdu word is derived from Persian whose

		source is the Avestan "vidava" and the French one is from the Latin "vidua". The root etymology is the Indo-European *widhewa, which is the feminine of an adjective *weidh- meaning separate.
Vivre	جینا	The Sanskrit word "jiva" is the source of our Urdu word while the Latin word "vivere" is the source of the French word. Proto-Italic***gwiwe/o-*** 'to live', ***gwiwo-*** 'alive', ***gwIta-***.They all in turn derive from the Indo-European *gwih3-ue/o- 'to live', *gwih3-uo- 'alive'.
Voix	اواز	The Urdu word is derived from Persian whose source is the Old Avestan "vax" and the French one is from the Latin "vox". The root etymology is the Indo-European *uok-s.

Conclusion

Urdu and French languages have shown both linguistic and historical similarities. The level of transparency of relationship exhibited between French and Urdu is not intuitively evident for specialists and native speakers alike. In fact, it is difficult to discern that these languages have arisen from a common genetic source unlike in the case with Romance, Germanic or Turkic languages. Therefore, special research is needed to demonstrate it by comparing such lexico-syntactic similarities that cannot be explained by any other linguistic relationship except genetic relatedness. As in this article it is proved that, there are certain observable similarities between Urdu and French in the areas of phonology, morphology, syntax and lexicon because of their common genetic connection.

The history of the development of Urdu parallels with the history of the French Language. Just as French evolved from a dead language of Europe (Latin) Urdu also evolved from the dead language of India (Sanskrit). They emerged as "vernaculars" of their classical languages (Latin and Sanskrit). Both developments also occurred almost simultaneously during the last millennium. Another similarity is that both languages (Urdu and French) in their lexical development borrowed words from other languages. In the case of Urdu, the contribution is from Arabic and Persian and in the case of French from Germanic, Celtic and other Romance languages.

Although the languages have very different scripts, Urdu and French share very similar sounds - with the result that, for speakers of Urdu, French phonology does not create any problem, as almost all French sounds are part of the Urdu phonology. However, the presence of certain sounds in Urdu such as aspiration and retroflexion makes Urdu difficult to pronounce for French or any other European language speaker.

END NOTES

i   The Constitution of Jammu and Kashmir, 2003

ii   Schedule VIII languages are those languages which are given in the section VIII of the Indian Constitution and the government of India is under an obligation to take measures for the development of these languages.

iii   Organisation Internationale de la Francophonie, 2014

iv   http://andromeda.rutgers.edu/~jlynch/language.html

v   The following etymologies were extracted from "A Dictionary of Selected Synonyms in the Principal Indo-European Languages" by Charles D. Buck; "Etymological Dictionary of Latin and Other Allied Languages" by Micjiel De Vaan; "The American Heritage Dictionary of Indo-European Roots" by Calvert Watkins and " Urdu Encyclopedia Dictionary" (consulted online). Collins Robert French Dictionary is consulted for French pronunciation.

References

Aitchison, J. (2001). *Language Change: Progress or Decay?* Cambridge: Cambridge University Press. 6

Boyd-Bowman, P. (1980). *From Latin to Romance in Sound Charts.* Washington: Georgetown University Press. 50-107.

Brown, K. & Ogilvie, S. (2009). *Concise Encyclopedia of Languages of the World*. Oxford: Elsevier Ltd. 1133-1135.

Buck, Ch. D. (1949). *A Dictionary of Selected Synonyms in the Principal Indo-European Languages.* Chicago: University of Chicago Press.

Campbell, L. & Poser J. W. (2008). *Language Classification*. Cambridge: Cambridge University Press. 165.

Collins R. (2010). *French Dictionary*. UK: Collins Publishers.

Comrie, B. (2009). *The World's Major Languages.* London: Routledge. 171,400.

Clackson, J. (2007). *Indo-European Linguistics.* Cambridge: Cambridge University Press.

De Vaan, M. (2008). *Etymological Dictionary of Latin and the other Italic Languages*. Leiden: Brill Publishers.

Desai, U. & Ramsay-Brijball, M. (2004). Tracing Gujarati Language Development Philologically and Linguistically. *AlterNation : Journal of the Centre for the Study of Southern African Literature and Languages*. 11 (2), 310-312.

Emeneau M.B. (2000) *Sanskrit.* Encyclopedia Americana. 24. 232-233.

Fagyal, Z., Kibbee, D. & Jenkins, F. (2006). *French: A Linguistic Introduction*. Cambridge: Cambridge University Press. 5.

Gell-Mann, M., Peiros, I. & Starostin, G. (2009). Distant Language Relationship: The Current Perspective. *Journal of Language Relationship*s, 1(1), 13.

Joseph, B. D. and Janda, R. D. (Eds.) (2003). *The Handbook of Historical Linguistics*. Oxford: Blackwell. 183-186.

Khan M. H. (1969). 'Urdu.' In the book T A (ed.) *Current Trends in Linguistics*, V. The Hague: Mouton. 277–283.

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). (2016). *Ethnologue: Languages of the World,* Nineteenth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com

Lockwood, W. B. (1972). *A Panorama of Indo-European Languages*. (1-2) London: Hutchinson Press.

Marcucci, E. M. (1855). *Lettere Edite e Inedite di Pilippo Sassetti Raccolte e Annotate*. Firenze: Felice Le Monnier. 417.

Mukherjee, B. N. (1997-98). The Unsolved Aryan Problem. Indologica Taurinensia. Vol. 23-24. 72-73.

Organisation Internationale de la Francophonie. (2014). *La Langue Française dans le Monde*. Paris. 3-4.: http://www.francophonie.org/IMG/pdf/oif_synthese_francais.pdf (Retrieved on 5th September, 2015).

Swadesh, Morris. (1952). Lexicostatistic Dating of Prehistoric Ethnic Contacts. *Proceedings of the American Philosophical Society,* 96, 452–463.

Swadesh, Morris. (1971). *The Origin and Diversification of Language*. Edited post mortem by Joel Sherzer. Chicago: Aldine.

The Census of India, 2001. Retrieved from http://www.censusindia.gov.in/Census_Data_2001/Census_Data_Online/Language/Statement4.aspx (Retrieved on 4th September, 2015)

The Census of Pakistan, (1998). Population by Mother Tongue: http://www.pbs.gov.pk/sites/default/files//tables/POPULATION%20BY%20MOTHER20TONGUE.pdf (Retrieved on 11th July, 2015)

The Constitution of Jammu and Kashmir. (2003). P. 41 http://jklegislativeassembly.nic.in/Costitution_of_J&K.pdf (Retrieved on 6th September, 2015)

Urdu Encyclopedia Dictionary (Online Version). http://www.urduencyclopedia.org/urdudictionary/

Walter, H. (1994) French Inside Out. (Translated by Peter Fawcett). Routledge. London. 129

Watkins, C. W. (2000). The American Heritage dictionary of Indo-European Roots. Boston: Houghton Mifflin Co.