Linguistics and Literature Review (LLR)

_____

Linguistic Simplicity and Complexity in Computer- Mediated Communication of Urdu/English Bilinguals

Dr. Muhammad Shaban Rafi

Article QR Code:

Shaban R

A publication of the
Department of English Language and Literature
School of Social Sciences and Humanities
University of Management and Technology
Lahore, Pakistan

Linguistic Simplicity and Complexity in Computer-mediated Communication of Urdu/English Bilinguals

Dr. Muhammad Shaban Rafi

University of Management and Technology – Lahore, Pakistan

ABSTRACT

This study examines how linguistic practices of Urdu/English bilinguals influence linguistic typology particularly in terms of linguistic simplicity and complexity. The data was sampled from the Bachelor of Science students (who had Urdu as their primary language of communication and English as one of the academic languages or the most prestigious second language) of five universities located in Lahore, Pakistan. The data was primarily from their Facebook communication on the wall. The procedure for analysis was conceived within the current theoretical work on text analysis. At any given moment in time, interpersonal communication of Urdu/English bilinguals shows linguistic simplicity and complexity. The linguistic features which involve complexity are generally avoided and linguistic simplicity is emerging as the norm. The diachronic analysis of the data supports non-complexity axiom and further shows that the linguistic variations which used to occur over a period of decades are presumably spreading in a matter of years.

Introduction

Linguistic simplicity and complexity can be viewed as an outcome of contact between speakers of different languages (e.g., Trudgill 2011). Linguistic simplification can be defined through certain linguistic features such as an increase in lexical and morphological transparency, loss of allomorphy and regularization of irregularities. On the other hand, linguistic complexity is defined as a diametrically opposite perspective that involves an increase in linguistic opacity, increase in linguistic redundancy, and irregularization of regular forms. Going back to the concept of language contact that takes place in bilingual or multilingual contexts where there is an increased social contact between speakers who have traditionally spoken different languages.

Language contact or strictly speaking contact between speakers is usually motivated by various factors such as economy, education, tourism, colonization and so on (Rafi 2013: 1260-1275). Moreover, we may not overlook the phenomenon of globalisation, which is pulling people together across space and time. Ever increasing contact between speakers of diverse social and linguistic background can be ascribed to Internet based communication. It is important to note that contact between different languages in the digital environment is presumably a little more complicated. Ever increasing digitalization among bilinguals/multilinguals seems to evoke particular types of language structures, which are, of course, different from what we have been discussing as a result of language contact for several years (c.f. Fischer 1958; Labov 1966; Trudgill 1974; Milroy 1980; Llamas 2007).

The proliferation of Internet-based communication has provided an opportunity to bilingual speakers to invent unique linguistic forms and to do frequent code-switching. Trudgill (2003) defines bilingualism as "the ability of an individual to speak two or more languages" (15). Rafi (2014: 2) argues that the use of Urdu on the Internet is always embedded in the larger Anglophone context. Urdu/English bilinguals draw on diverse linguistic resources which not only motivate new patterns but also limit them. Depending upon the topic and situation, Urdu/English bilinguals usually juxtapose their linguistic knowledge of at least two languages at word, phrase and clause level (Rafi 2013: 1272). This fact is compatible with Aure's (1999) continuum of language alternation e.g., code-switching, language alternation and fused lects. While distinguishing these terminologies Peuronen (2008: 16) explains that code-switching covers the pragmatic pole and fused lects define the grammatical pole on the continuum, however language alternation falls between these poles. Gumperz (1982) introduced the concept of *we-code* and *they-code* to explain how translanguaging practice reflects interculturality as noted by Hua (2015: 109-124). Bilinguals choose to oscillate freely between available language structures while code-switching (e.g., Muysken 2000; Demircay and Backus 2014). Many researchers (Baron 2008; Bodomo 2009; Crystal 2006; Kim, Weber, Wei and Oh 2014; Rafi 2014 and 2017) assert that bilingualism/multilingualism has now become a norm in CMC.

As CMC has been increasingly becoming multilingual, researchers (e.g., Androutsopoulos 2007; Axelsson, Abelin and Schroeder 2007; Barasa 2010; Bodomo 2009; Durham 2007; Leppänen 2007; Paolillo 2007; Seargeant, Tagg and Ngampramuan 2012; Rafi 2014 and 2017; Warschauer, El Said and Zohry 2007) have explored new patterns of use and language combination in bilingual/multilingual speech communities.The linguistic repertoire of Urdu/English bilinguals involves frequent switches, linguistic reduction and neologisms. Crystal (2006) argues that "Internet users are continually searching for words to describe their experiences, to capture the character of electronic world, and to overcome the communicative limitations of its technology" (67). The prevalence of certain linguistic forms, both in their linguistic choices and elsewhere, seems to construct social relationships within the global context (c.f. Wei and Hua 2013: 516-535).

Part of the Internet sociology lies in what is getting simplified and/or complexified in the linguistic repertoire of Urdu/English bilinguals in CMC. Throughout the history, we have

witnessed linguistic simplification at a gradual and slow pace. But it seems that CMC is accelerating the speed of linguistic simplification. Unlike any other study of language variation and change, cyber-linguistics allows us to follow, like never, the rate and reach of language variation (Crystal 2006). Arguably, new-media language research has encouraged more scholars to investigate the scale of impact of CMC on languages used widely on the Internet. In contrast with the past studies (c.f., Trudgill, 2011 and those cited therein) which have noted striking differences in the speed of linguistic simplification and complexification at phonological, lexical, syntactical and semantic levels of different languages and dialects, the present study hypothesises that Internet communication has influenced crucially the speed of linguistic variation in the Urdu language.

Material and Methods

Data Collection

Data collection was delimited to Facebook wall, which gave an easy access to an enormous pool of data, without violating anyone's privacy. The study considered communication on Facebook wall analogous to asynchronous communication between one-to-one and one-to-many. Interactivity is a defining characteristic of Facebook conversation on the wall, which is organized around various topics that interest to users. Further, it can be characterized by the use of a broad range of traits at the very informal end of the linguistic spectrum, which in a way is comparable with face-to-face conversation between the users.

Sample

Five volunteers who were students at the sampled institutions were engaged for assistance in the process of data collection. They were aware of the linguistic and cultural background of the selected participants. Each volunteer coordinated the data gathering process for one of the five institutions. The researcher shared with them the purpose and ethical boundaries of the study. The same things were shared with the participants as well. Each volunteer successfully added on average 375 students over a period of two months. The researcher had access to all the students through these volunteers. The data from 50 participants from each institution was analysed on the basis of the frequency of their status updates. The study investigated linguistic postings of each participant transmitted over a week. The reason for collecting the whole week's data was to observe maximum number of linguistic features that the participants used in their communication. Each posting consisted of usually more than one utterance. Hence, the data collected was naturalistic and observational.

Hence, the data was sampled from 200 male and female students who were between 18-24 years of age and were registered in Bachelor of Science (BS) program in the five private (by and large not funded by the State) institutions of higher learning. The reason to choose these institutions was to make the sample as representative of BS students as possible. In terms of academic, linguistic and cultural background the sample was homogenous. All the participants were from BS program but they were from different disciplines (e.g., Business of

Administration, Engineering, Computer Sciences and English) and academic years (e.g., covering first year to fourth year). They were between 18-24 years of age. More importantly, their communication on Facebook was more or less with the users of same age group. Notwithstanding the possibility of this forum to connect people from different linguistic backgrounds, the participants in this study were evidently Urdu/English bilinguals. It is important to note that Urdu was their primary language of communication, both at home and in various other settings, however English was their most important academic language and the most prestigious second language. The topics of their conversations were not only limited to the Pakistani context but expanded to include the international context as well.

Nature of the data

The postings on the wall grew around topics of varying kinds. Mostly, a topic was initiated by a single person followed with his or her followers' comments, which were either closed with words of gratitude or reflection by the originator or left without proper closing. The exchanges on the wall can be classified into a three-part structure of Initiation, Response and Reflection (IRR), dialogue, two-pair adjacency and non-linear conversation. A linguistic posting on the wall can be generally described as consisting of fragments, phrases, clauses and short paragraphs. A lot of words that the participants used were clearly from their specific academic context, however the tone used in the message threads was both formal and informal. On the other hand; they carried out various activities e.g., information exchange, debate, problem solving, exchanging picture, video and jokes, which can be grouped under various themes such as greeting, politics, religion, showbiz, sports, education and sex. The data was mainly in romanized Urdu, English, or a mixture of romanized Urdu and English. The participants used largely, if not completely, roman script in their status updates. Probably, the relative ease of typing in Roman vs. Urdu script give the computer setup/keyboard common for the participants.

Ethics of data collection and handling

The ethical guidelines suggested by Mann and Stewart (2000: 40-47) were followed for collecting and handling the data. The participants were informed about the nature of the study. They were given assurances regarding confidentiality, security of information, and unauthorized eavesdropping; that is, information that might identify their place, institution and time was never to be disclosed. They were masked by means of cryptonyms in reporting the findings. Access to the data was restricted to the researcher and the volunteers. Since the data was collected in the context of free conversation, the researcher could not forbid the use of racist and sexist language, and other contentious and provocative material.

Data Analysis

We cannot simply speculate linguistic simplicity or/and complexity as an outcome of contact between languages only. There seems to be a good case for looking for forms getting simplified and/or complexified in the linguistic repertoire of Urdu/English bilinguals. The linguistic simplicity and complexity was examined through lexical and syntactic features. I have adopted

the approach outlined in Trudgill (2011) which explains simplification in terms of (a) the regularization of irregularities (b) an increase in lexical and morphological transparency and (c) the loss of redundancy. The notion of complexity was studied by reversing these processes (a-c). Trudgill (2011) argues that "if simplification can be characterised in terms of these processes then it follows that complexification consists of the reverse processes" (62). I have relied on material from Urdu and English that provided a suitable context to explore what is getting simplified or/and complexified in the linguistic repertoire of Urdu/English bilinguals.

Moreover, the analysis of the participants' timelines has allowed me to observe the temporal factor in the discussion on linguistic simplicity and complexity. This will provide us a resource to address the assumption regarding linguistic variation that used to happen over a long period of time but it is now spreading in a comparatively short time period. Frequency was calculated to gauge how many times a particular feature occurred and to suggest its permanence in the conversation of Urdu/English bilinguals which provided a clue about the potential of a feature to reside in the system of the recipient language. The frequency of occurrence was determined if a word was repeated twice within a conversation and a minimum of five times in the whole data. As many as 2,516 linguistic postings were studied to address the research question. Of these postings, 588 were in Romanized Urdu; 1135 were in English; and remaining postings were a blend of both Urdu and English. On average, each posting consisted of 137 words.

In addition to the analysis of quantitative data, I drew on the message threads to elaborate and support my verdict concerning the research question, which helped bring triangulation to increase the credibility and validity in the results. These snippets were demonstrated by mathematical symbols (such as <>) along with their transliteration in the parenthesis. In relation to paralinguistic features, I considered their visual aspects only as a meaning making resource in a message thread. Thus, the analysis was backed up by a large corpus covering both quantitative and qualitative data sets, which assisted in answering the research question with more extensive analysis, with far more participants and much more rigorous sampling procedures.

Simplification

The data shows that Urdu has been going through a kind of contact with other languages that leads to linguistic simplicity. Linguistic reduction (e.g., logogram, abbreviation, clipping and orthographic reduction) and neologism (e.g., morphemic substitution, coinage and derivation) are features comparable with morphological transparency. The most frequently used feature of text orthography – the use of phonetic spelling, lexo-numeric, digito- lexeme and digit word homophone – is logogram. Phonetic spellings are sounds which are occasionally shortened to mono-syllabic or bi‑syllabic sounds, e.g., *u* for you, *r* for are, *y* for why and so on. The participants create lexo-numeric words by compounding a number and lexeme or morpheme, e.g., *some1* for someone, *gr8* for great, *b4* for before and so on. It is evident from the examples that lexo-numeric words are composed of two syllables in which one syllable is substituted with a digit and the other remains constant. Conversely, when we shuffle the sequence of lexo-numerics we come across another category that we may label digito-lexeme. This category

involves the replacement of segment or segments of a word with a digit, e.g., *2morow* for tomorrow, *4get* for forget and so on.

Among other categories of reduction, digit word homophones are prevalent, which are formed by replacing a full word with a digit, e.g., *1* for one, *2* for two or too, *4* for for and so on. These examples evidently mirror a prevalence of morphological transparency.

The graphemic approximation in Urdu (e.g., ن for /r/, ر، ڑ، ڈ، د for/s/, ث، س، ص for/a/, ا،ع for/a/, ص، س، ث for/s/, ز، ژ، ض،ظ for /z/, ط، ت، ٹ for /t/, ق، ک for /k or q/, غ، گ for /g/ and ه، ح،ہ for /h/) instantiates an increase in morphemic transparency as noted by Rafi (2014: 7). There are thirty-eight alphabets in Urdu but in romanized Urdu 66% of them are reduced to 24%. Around 58% of Urdu alphabets find seemingly semi-homophonous corresponding letters in English. As a result of this, Urdu multi-syllabic words are replaced with mono-syllabic forms in English. The participants mostly reduce English forms, however they simply substitute bi-syllabic Urdu words with the whole corresponding mono-syllabic English homophones or semi-homophones – may be considered a creative way of reducing morphemic properties of Urdu primarily in romanized script. While substituting Urdu forms with its English counterparts, they overextend the sounds have less homophonous correspondence in some instances, e.g., بھی /bʰi/ with [b]. It is evident that the substitution does not cover aspirated sound because this requires use of superscript [ʰ] that the participants simply avoid.

[b], pronounced as /bʰi/is a semi-homophone of بھی which means 'also'

[c], pronounced as /si/is a homophone of سی which means 'also or 'like' or 'of' or 'for'

[i], pronounced as /ai/ is a homophone of ائی which means 'coming'

[g], pronounced as /dʒiː/is a homophone of جی which means 'yes'

[k], pronounced as /keɪ/ is a homophone of کے which means 'that' [q], pronounced as /kju/ is a homophone کیوں which means 'why'

The reason behind these substitutions is that all the above-mentioned Urdu bi-syllabic words have the same or nearly the same sound in English alphabets. It seems, more likely, as a matter of ease that eventually causes morphological transparency since the participants replace bi-syllabic forms with their mono-syllabic counterparts. Similarly, they also substitute the following English words with their counterparts in Urdu.

[a], is pronounced as /ɑː/ is a homophone of آ which means 'come'	[may], is pronounced as /meɪ/ is a homophone of میں which means 'in'
[gay], is pronounced as /geɪ/ is a homophone of گے which 'went' or 'past forms of will (would)'	[or], is pronounced as /ɔː(r)/ is a homophone of اور which means 'more')
[he/hi], is pronounced as /hiː/ is a homophone of ہی which is 'intensifier'	[pass], is pronounced as /pæs/ is a homophone of پاس which means 'near'
[her], is pronounced as /hɜː(r)/ is a homophone of ہر which means 'every'	[pay], is pronounced as /peɪ/ is a homophone of پے which means 'on'
[key], is pronounced as /kiː/ is a homophone of کی which means 'what'	[say], is pronounced as /seɪ/ is a homophone of سے which means 'from'

[log], is pronounced as /lɒg/ is a homophone of لوگ which means 'people'	[such], is pronounced as /sʌtʃ/ is a homophone of سچ which means 'true'

The second most noticeable feature of morphological transparency is the reduction of words to their initial letters. They are known as abbreviations. Abbreviations are commonly used and there is an accepted trend of writing the initials of words to make a new word that is not essentially pronounceable. However, there are instances of abbreviations which obey phonetic properties, e.g., *AFAP* for as far as possible, *SOB* for son of bitch, *Yo* for your own and so on. Abbreviations are usually spelled in capital letters, however there are cases where strings of words are abbreviated in lower-case letters also. Like reduction in base forms, abbreviations also involve the loss of material. The principle of orthography is, however of central importance in abbreviations, e.g., *dp* for digital picture, *gf* for girlfriend, *np* no problem and so on. As mentioned above, in limited cases, phonetic properties are also applied to derive abbreviations known as acronyms, e.g., *asap* for as soon as possible, *lol* for lots of laughter, *afap* for as far as possible and so on.

There are words which have been derived from the first part of the base word which can be labelled as clipping and it also supports the process of morphological transparency. Clippings appear as a mixed bag of forms reduced from base forms, which expresses familiarity with the denotation of the derivative. Thus, *pic* is used typically by the participants to refer to digital image and *bro* as part of their vocabulary to show probably an intimate relationship. In the following list, there are clippings, e.g., *add/addy,* cos, *del*, *dif*, *grats*, *pic*, *prop*, *min*, *morn*, *rem*, *uni*, and *web* which are characteristic features of morphological transparency.

Along with the minimal linguistic forms (c.f. Rafi 2014); neologism, though less prevalent as compared to linguistic reduction, mirror morphological transparency. One of the newly devised categories of word-formation is the replacement of 's' with 'x' and 'z' sounds (where these transposed variants of *s* might be the equivalent sound at the end of lots), e.g., *cheerx* for cheers, *Lolz* for *lols* (lots of laughters), *lolx* for *lols*, *returnx* for returns, *shex* for she's, *thtx* for that's, *yupx* for yups (yes), and so on. However, in most of the instances, it can be observed that 's' is replaced with 'z' to perceive its natural sound. Contrary to the typical use of 'x' and 'z', mostly if not always, as an inflectional morpheme or suffix, there are a few instances where use of 's' and 'z' is overextended e.g., *Pathanz* (a speech community that lives in the North of Pakistan), *Kal's* (The expression stands for an accentuated either yesterday or tomorrow – only context discriminates the right option) in the Urdu language. Baumgardner, Kennedy and Shamim (1993: 143) reveal similar finding and explain how Urdu borrowings undergo the morphological rules of English retained in Pakistani English. They further remark that the combination of English affixes with Urdu has produced vocabulary previously unattested in other varieties of English. The use of an English suffix 's' to mark plurality by attaching it to Urdu words involves a kind of morphological transparency.

In addition to morphemic substitution, a small set of words is formulated by applying a rule of Urdu pluralisation with the English root, e.g., *Blockkan* for Blocks, *Copian* for Copies, *Filmma* for Films, *Simmon* for Sims, *Teacheran* for Teachers etc. In these examples, the root

words are borrowed from English but inflect for plurality by applying the rule available in Urdu. Though Urdu/English bilingual usually embed the English language into the structure of Urdu, CMC apparently provides a context to ungrammatical formation as noted above. Baumgardner Kennedy and Shamim (1993: 43) argue that neologism is an area in which Pakistani English has forged its own identity.

Unlike the above mentioned categories of words-formation, only a single token of derivation is found concurrently in the data. A derivation is accomplished by means of the combination of a large number of small bits, such as affixes, which are not usually assigned separate listings in dictionaries. The data surface an instance of @ for 'at' prefix e.g., *@cafe* for at café, *@home* for at home, and @uni for at university and so on. The use of this sign before a proper noun e.g., @zeshan to address a participant in the message thread is not uncommon in CMC. Though least common, this feature is quite prominent that suggests the linguistic simplicity.

Unlike morphological transparency, there are instances of structural economy e.g., marking of mood, tense and voice without the structural transformation that warrants linguistic simplicity. In attempting to express themselves, the participants do not only exchange simple utterances, they also employ different clausal moods. For instance; in [1a] imperative clause structure is used along with the insertion of exclamatory tone to enquire. Similarly in [1b and 1c] declarative structure is used to enquire. In [1c] use of a question mark at the boundary indicates that the participant has an intention to enquire. In [1d] the structure is interrogative but the exclamation mark and comma seems to indicate the intention. Thus, the structure alone may not be sufficient to perceive the communicative function in CMC. How the participants use their linguistic repertoire to perform syntactic expression of mood seemingly depends upon their common understanding of the unique structures.

[1]

    a. *<Tell me abt it!>* (Tell me about it!)
    b.*<Mje call pe btana Kal paper hai mera.>* (Let me know on phone … I have exam tomorrow.)
    c. *<Kal match on?>*(.. .. … game on tomorrow?)
    d. *<Phr kon dayta hay apko, clases,,?!>*(Then, who does teach you?)

The deletion of auxiliaries, clauses (e.g., finite, infinite and non-finite), prepositions and articles support structural transparency. The participants occasionally truncate 'I am' to 'am' or 'm' or simply omitted 'I' e.g.,*… will do asap… m having my exams!!... hope u understand!! … love you mwaaaah*. Similarly, they omit auxiliary verbs, such as is, are, am, was and were in their conversations. Given the structural deletion, there is omission of capitalisation in the beginning of an utterance or in the case of proper nouns and addition of toggle case. As a result of this, communication is more a reflection of short structures then supported by linguistic rules. The omission of linguistic features shows that the participants may have used their pragmatic knowledge to presuppose that the receiver knows how to map out deleted expressions in their utterances.

Apart from obligatory words, the participants omit optional words by assuming that obligatory words might be sufficient to express meanings. On the other hand, the English language resists omission of obligatory strings of words because they bear meanings and the omission may damage intelligibility. In my observation, the deictic expressions; main verb; and attributive forms seem sufficient for the projection of deleted strings of words. There is reason to believe that even though their communication is structurally simplified, concurrent exchanges reveal somehow that the participants may have inferred the deleted string of words for the success of communication. It is logical to argue that meaning limits the structure, which is analogous to non-complexity axiom (e.g., Trudgill 2011: 20-26).

Corresponding to linguistic reduction that evidently indicates simplification; another process that matters is almost no use of aspirated sound /^{h}/ (e.g., ھ ،ہ ،ح) in the romanized Urdu at the word medial or ending position which supports the second process – loss of allomorphy. Kusters (2003) argues that loss of allomorphy is comparable with lexical and morphological transparency. There are several words in Urdu which are typically used without an aspirated sound; some of them are enlisted in [2]. One may assume that such sound involves complex articulation which obscures its production in a speech but the aspirated free orthography leads us to the finding that the participants seem to be simplifying phonology in the digital discourse (e.g., Azim 2002: 273-307). There is a fair chance that phonology being a cognitive trait may remain unchanged except when motivated by its orthographic form. It may be a little deceptive to draw a straightforward conclusion regarding   the relation between phonology and orthography. But we can surely say Urdu is undergoing morphological simplification.

[2]

[ri] is used for /ɪhi/ which means 'also'
[tmary], is used for /tʊmhaɾi/ which means 'your'
[chaiya], is used for /t͡ʃahɛ/ which means [masla],
is used for /mǝslǝ/ which means 'issue' [b], is
used for / bʰi/ which means 'also'
[ta], is used for /ɵǝ/ which means 'was'

The data shows instances of voiced nasal alveolar /ʊ / which is realised though /n/ is dropped at the word final position. As shown in [3], /ʊ/ was mostly if not always dropped, seemingly to simplify the orthography. Probably, this behavior can only be generalized in the digital discourse.

[3]

[karu], is used for /kǝrʊn/ which means 'doing'
[ha], is used for /ɦein/ which means 'is or are'
[ni or nai or nae], is used for /nɦin/ which means 'not or no'

As noted above, the linguistic repertoire of the Urdu/English bilinguals includes simplification in terms of an increase in lexical and morphological transparency and loss of allomorph. These processes are pervasive throughout the data and they support the third process that involves the regularization of irregularities. The examples we have noted above to support a proliferation of

the first two processes seem to motivate regularization of irregularities as opposed to the use of conventional forms.

Complexification

Let's now examine a diametrically opposed standpoint that involves: (a) increase in opacity, (b) increase in redundancy, and (c) irregularization. What follows is a brief account of these processes. I have drawn on not many examples of each of these processes, however the instances which are presented clearly support the complexity axiom. The data shows forms (e.g., *czn*, *bzzz*, *gr* and so on) which are opaque. They may lead us to different interpretation until we contextualize them for some plausible guess. For example, *czn* can be read as 'season', bzzz may lead us to interpret it as 'buzz' and *gr* may be thought out as grass, grow and gross, whereas, the context help us to figure out them as 'cousin', 'bus' and /gʰr/ which means 'house'. Though irregular, these forms are more or less idiosyncratic which may cause ambiguity unless they become regular or perceived within the appropriate context. So, we can say that part of linguistic complexity is a source of pragmatic obscurity. The only instance that we find to correlate with the process of redundancy is repetition of some letters (e.g., *happyyyy* for happy, *thnxxxx* for thanks, *soooo* gorgeous for so gorgeous) probably to give an impression of verbosity. Dahl (2004) defines verbosity as containing more material than would be necessary (as cited in Trudgill 2011: 53). The data does not mark instances of linguistic redundancy, the reason maybe the participants chose to economize their linguistic forms – a characteristic feature of linguistic simplicity. On the other hand, the processes we have examined such as increasing opacity, morphemic addition and irregularization lead us to expect complexity-development.

The examples noted above verify that both linguistic simplicity and complexity are occurring side by side in the linguistic repertoire of Urdu/English bilinguals. The figure 1 summarizes what happens when Urdu/English bilinguals have their conversation in CMC. Moreover, diachronic variation in the data provides further a resource to view the temporal factor in the processes of simplification and comeplexification. The diachronic perspective that I have taken can also be used as a correlate of the speed of language variation and change.
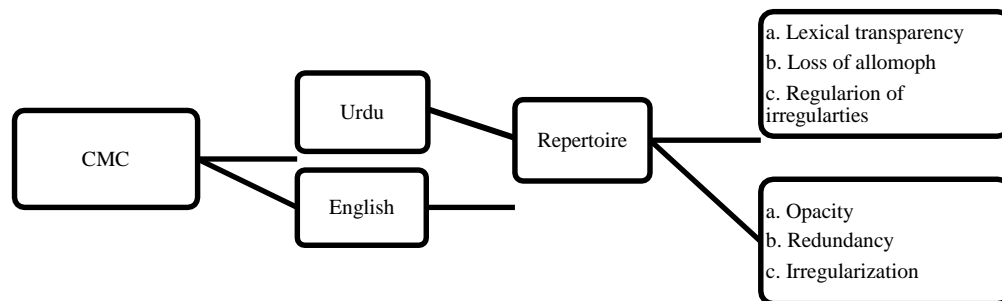
Figure 1. Simplification and complexification in the linguistic repertoire of Urdu/English bilingualsin CMC

Diachronic Variation

The diachronic variation demonstrated in table 1 has been slow but gradual for over four years. Note that the sign (+) shows the strength or high frequency of the corresponding feature; whereas the sign (-) indicates weakness or low frequency of the feature, however co-occurrence of the signs (+ and -) symbolises relatively less strength and co-occurrence of the signs (- and +) indexes the least common feature in the data. The table 1 indicates that mixing of English words has been frequent as compared to the mixing of phrases and clauses between 2009 and 2012. The data did not record a single instance of mixing of English clause until 2009, however it has been growing ever since. It is important to note that ever increasing frequency of English words into Urdu seems to be blurring the distinctions between matrix language and second or foreign language. The base language in which the second or foreign language is mixed is called matrix language (c.f. Myers-Scotton 1998). The current situation that deserves our attention highlights the lack of or non-existence of planning in Pakistan that naturally allows English to rule over the Urdu language.

Table 1. Diachronic variation in the linguistic repertoire of Urdu/English bilinguals

Year	Language mixing			Linguistic reduction		Neologism
	Word	Phrase	Clause	Word	Structure	
2012	+	+	+	+	+	+
2011	+	+	+-	+	+	+
2010	+	+-	-+	+	+	+
2009	+	-+	-	+	+-	-

The diachronic analysis further underpins the fact that lexical reduction has been quite pervasive, however structural reduction was less common in 2009. In addition to this, dropping of 'g','t' and 'd' (e.g., *talkin* for talking, *lil* for little and *n* for and) in the data points out the spread of native-English pronunciation that attributes language change in the linguistic inventory of the participants. I doubt if such linguistic forms were part of the linguistic resource of young Pakistani students until the popularity of CMC in their lives. It is perhaps hard to predict when and how these linguistic features migrated but they unarguably show an influence of native-English on the linguistic choices of the participants. They regularly minimized lexical and syntactic forms of English, however the only instances of reduction that I discovered in Urdu are the replacement of its graphemes with English phonemes and substitutions of Urdu bi-syllabic words with English mono-syllabic forms. This finding underlines the use of common Urdu and English sounds, however Urdu phonology is overgeneralized while simplifying morphemic or/and structural properties

New words are thought to be ideational. They are formulated to refer to ideas which were previously never expressed by other words or in some cases might not have existed. The present study shows a decreasing/marginal need for new words until 2009. The table shows they have been proliferating quite frequently ever since. Most of them are idiosyncratic embellishments, however some are newer and more stylish ways of creating words to express meanings.

Conclusion

The linguistic repertoire of Urdu/English bilinguals in CMC provides an insight about simplification and complexification processes. The study deviates slightly from Trudgill's hypothesis (2011) that high-contact varieties of English are characterised by structural simplification processes while low-contact varieties are the result of complexification processes. This study, on the other hand, demonstrates that linguistic simplicity and complexity established in the data go hand- in-hand, irrespective of the nature of contact, in the creolised variety of Urdu and English. This finding aligns supposedly with the results of Kortmann and Szmrecsanyi (2009), and Chand (2012).We can conclude that linguistic consequences which used to happen over a long-time can now be viewed to a considerable extent in a short time period in the context of CMC. The study suggests drawing an independent investigation on pragmatic and semantic consequences as a result of morphological and structural changes observed in the linguistic repertoire of Urdu/English bilinguals.

References

Androutsopoulos, J. 2007. Language choice and code switching in German-based diasporic web forums. In *The Multilingual Internet: Language, Culture, and Communication Online, eds.* B. Danet and S. Herring, 340-361. Oxford: Oxford University Press.

Auer, P. 1999. From code-switching via language mixing to fused lects: Towards a dynamic typology of bilingual speech. *The International Journal of Bilingualism* 3(4): 309-332.

Axelsson, A.S., Abelin, A., and Schroeder, R. 2007. Anyone speak Swedish? Tolerance for language shift in graphical multiuser virtual environments. In*The Multilingual Internet: Language, Culture, and Communication Online, eds.* B. Danet and S. Herring, 362-381. Oxford: Oxford University Press.

Azim. A. 2002. Problems of aspiration in modern standard Urdu. In *Signal, Meaning, and Message: Perspectives on Sign-based Linguistics*,eds. H. W.Reid, R.Otheguy and N Stern, 273-307. Amsterdam: John Benjamins Publishing Company.

Barasa, S. 2010. *Language, mobile phones and Internet: A study of SMS texting, email, IM and SNS Chats in computer mediated communication (CMC) in Kenya*. Janskerkhof: LOT publisher.

Baron, N. S. 2008. *Always on: Language in an online and mobile world.* New York: Oxford University Press.

Baumgardner, R. J., Kennedy, A. E. H., and Shamim, F. 1993. The Urduisation of English in Pakistan. In*The English Language in Pakistan*, ed. R. J.Baumgardner, 83-203. Karachi: Oxford University Press.

Bodomo, A. B. 2009. *Computer-mediated communication for linguistics and literacy: Technology and natural language education*. Hershey: IGI Global.

Boyd, D., and Ellison, N. 2007. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication* 13(1): 10-25.

Chand, V. 2012, October 10. Direction of change in urban Indian English. Seminar Talk at the School of English, Communication and Philosophy, Cardiff University, Wales, UK.

Crystal, D. 2006. *Language and the internet*. Cambridge: Cambridge University Press.

Demircay, D., and Backus, A. M. 2014. Bilingual constructions: Reassessing the typology of code-switching. *Dutch Journal of Applied Linguistics* 3(1): 30-44.

Durham, M. 2007. Language choice on a Swiss mailing list. In *The Multilingual Internet: Language, Culture, and Communication Online, eds.* B. Danet and S. Herring, 319-339. Oxford: Oxford University Press.

Fischer, L. J. 1958. Social Influence in the choice of a linguistic variant. In *An Introduction to Sociolinguistics*, 1990. R.Wardhaugh, 155-157. Massachusetts: Basil Blackwell Inc.

Gumperz, J. J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Herring, S. C., Stein, D., and Virtanen, T. 2013. Introduction to the pragmatics of computer-mediated communication. In *Handbook of Pragmatics of Computer-mediated Communication,*eds. S. Herring, D. Stein and T. Virtanen, 3-31. Berlin: Mouton De-Gruyter.

Hua, Z. 2015. Interculturality: reconceptualising cultural memberships and identities through translanguaging practice. In *Researching identity and interculturality. Routledge studies in language and intercultural communication*, eds., F.Dervin and K. Risager, 109-124. London: Routledge.

Kim, S., Weber, I., Wei, L., and Oh, A. 2014. Sociolinguistic analysis of Twitter in multilingual society. Proceedings of the 25th ACM conference on Hypertext and Social Media, 243-248.

Kortmann, B., and Szmrecsanyi, B. 2009. World Englishes between simplification and complexification. In *World Englishes-Problems, Properties and Prospects,*eds. T.Hoffmann and L. Siebers, 265-285. Amsterdam: John Benjamins Publishing Company.

Kusters, W. 2003. Linguistic complexity: The influence of social change on verbal inflection. In.*Sociolinguistic Typology: Social Determinants of Linguistic Complexity*, 2011. P.Trudgill, 21. Oxford: Oxford University Press.

Labov, W. 1966. The social stratification of English in New York City. In*An Introduction to Sociolinguistics,* 1990. R. Wardhaugh, 57-164. Massachusetts: Basil Blackwell Inc.

Leppänen, S. 2007. Youth language in media contexts: Insights into the functions of English in Finland. *World Englishes* 26(2):149-169.

Llamas, C. 2007. A place between places: Language and identities in a border town. *Language and Society* 36(4): 579-604.

Paolillo, J. 2007. How much multilingualism? Language diversity on the internet. In*The Multilingual Internet: Language, Culture, and Communication Online, eds.* B. Danet and S. Herring, 408-430. Oxford: Oxford University Press.

Mann, C., and Stewart, F. 2000. *Internet communication and qualitative research*. London, Thousand Oaks, New Delhi: Sage Publications.

Muysken, P. 2000. *Bilingual speech: A typology of code mixing*. Cambridge: Cambridge University Press.

Myers-Scotton, C. 1998. Code-switching. In *The handbook of sociolinguistics*, ed.F. Coulmas, 149-161. Oxford: Blackwell.

Peuronen, S. 2008. *Bilingual practices in an online community: Code-switching and language mixing in community and identity construction*. Retrieved from: www.godspeed.fi.https://jyx.jyu.fi/dspace/bitstream/handle/123456789/18390/urn_nbn_fi_jyu200805061436.pdf?sequence=1.

Rafi, M. S. 2017. Bilingualism and identity construction in the digital discourse. *Journal of Multicultural Discourses*, doi: 10.1080/17447143.2017.1342649

Rafi, M. S. 2014. Meaning making through miminal linguistic forms in computer-mediated communication. *Sage Open*. doi: 10.1177/2158244014535939

Rafi, M. S. 2013. Urdu and English in an e-discourse variation in the theme of linguistic hegemony. *European Academic Research* 6(1): 1260-1275.

Rafi, M. S. 2013. Urdu and English contact in an e-discourse: Changes and implications. *Gomal University Journal of Research* 3(1): 78-86.

Seargeant, P., Tagg, C., and Ngampramuan, W. 2012. Language choice and addressivity strategies in Thai-English social network interactions. *Journal of Sociolinguistics* 16(4): 510-531.

Stubbs, M. 1996. *Text and corpus analysis*. Oxford, Cambridge, Mass: Blackwell Publishers.

Trudgill, P. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

Trudgill, P. 2003. *A glossary of sociolingusitcs*. New York: Oxfords University Press.

Trudgill, P. 1974/1983. Sociolinguistics . In *Sociolinguistics*, 2003. A. R. Hudson, 40. Cambridge: Cambridge University Press.

Warschauer, M., Said, R. El. G., and Zohry, A. 2007. Language choice online: Globalization and identity in Egypt. In *The Multilingual Internet: Language, Culture, and Communication Online, eds.* B. Danet and S. Herring, 303-318. Oxford: Oxford University Press.

Wei, L., and Hua, Z. 2013. Translanguaging identities: Creating transnational space through flexible multilingual practices amongst Chinese university students in the UK. *Applied Linguistics* 34(5): 516-535.