

ELENA BASHIR

## Urdu and Linguistics: A Fraught But Evolving Relationship

### 1. Introduction

I WAS HONORED to be invited to give a talk at the Urdu Humanities Conference held in Madison, Wisconsin on 14 October 2010. However, when I thought about this, I wondered, “How can I present anything at a conference on Urdu humanities? I would be like a crow among the swans—a linguist among the literary scholars.” However, since I am a committed, card-carrying crow, with no pretensions to being a swan yet admiring their beauty, I took my life in my hands and proceeded. This estrangement that I have felt between the worlds of Urdu scholarship and of linguistics is the theme of this paper. I will begin by describing the disconnect I have perceived between Urdu studies and linguistics, discuss what I see as some reasons for it, and end with what seems to be a rapprochement or a new phase of this relationship.

Both “Urdu” and “linguistics” are recent terms. “Urdu” was not in use as the name of a language until the latter half of the eighteenth century (Faruqi 2001, 23),<sup>1</sup> the language which has become Urdu having previously been known by a variety of other names. Similarly, for “linguistics,” the term “linguistic” first appeared as a noun in the sense of “the science of languages” or “philology” in 1837, and its plural “linguistics” appeared in this sense first in 1855 (Onions 1955, 1148), and did not come into wider use as name for this discipline until the latter part of the twentieth century.<sup>2</sup> Therefore, this discussion will necessarily focus on

---

<sup>1</sup>Bailey (1939, 264) cites a couplet written in 1782 in which “Urdu” is used as the name of the language.

<sup>2</sup>When looking for the relationships between scholars who study language and the antecedents of Urdu, one must look for references to “grammarians” and

developments since the middle of the twentieth century.<sup>3</sup>

## 2. Disconnect between Urdu and Linguistics

### 2.1 In India

Despite the relatively early development of the discipline of modern linguistics in India, there has been relatively little work in that country specifically on Urdu. For example, in the journal *Indian Linguistics*, from Vol. 1 (1931) through Vol. 26 (1965) there were no articles with the word “Urdu” alone in the title. Articles including the term “Hindi-Urdu” appeared in Vol. 27 (1966), Vol. 36 (1975), Vol. 39 (1978), Vol. 49 (1988). One article on “Urdu” each appeared in Vol. 49 (1988), Vol. 54 (1993), and Vol. 56 (1995). Of these three, two were by a Muslim author and one by a European. Two early Urdu-language papers by C. M. Naim (1956, 1957) concerned Urdu phonology, but were not followed by further work in linguistics. The 30th All-India Conference of Linguists in 2008 (Linguistic Society of India 2008) included only two papers with “Urdu” in the title, Shukla (2008) and Mustafa (2008). Recently, judging by information this author has been able to find, it seems that work done in India on Urdu mostly concerns preparation of pedagogical materials such as bilingual dictionaries and textbooks, or the holding of teacher-training workshops. Some Ph.D. dissertations have been written but are, unfortunately, available only in the universities where they were produced and not accessible outside of India (e.g., Hasnain 1985).

Studies of Urdu varieties other than Modern Standard North Indian Urdu have received attention perhaps disproportionate to their number of speakers or position in the sociopolitical language hierarchy. Dakhini<sup>4</sup> (Urdu) has been the focus of a significant number of studies, perhaps because the difference of this dialect from the North Indian standard distances it from various religio-political issues and provides a scholar with

---

“philologists” and the variety of names by which the language now called “Urdu” was previously called, e.g. Hindvi, Rekhta, etc. The earliest academic Department of Linguistics in the U.S. was founded in the mid 1930s at the University of Chicago.

<sup>3</sup>Thus much important work on the Urdu language that was done in the early decades of the twentieth century, which would now fall in the category of linguistic research, will not be discussed here. This includes the numerous important works of T. Grahame Bailey (e.g., Bailey 1922, 1929, 1931a, 1931b, 1934).

<sup>4</sup>Spelled variously “Dakhani,” “Dakhini,” “Dakhani,” or “Dakhni,” depending, it seems, on whether the roman representation is taken from the Devanagari or from the Urdu spelling.

an area in which he can pursue his strictly linguistic interests. Also, as a minority variety, it has an inherent attractiveness for many linguists. An early work on Dakkhini Urdu phonetics is Qadri (1930); Schmidt (1981) deals with the phonology, morphology and history of Dakkhini and includes some texts. Khan and Mustafa (1984) deals with finite verbs, while Mustafa (2000) is a more recent comprehensive, descriptive grammar. Arora (1986) and Arora and Subbarao (1988, 1989) are studies of convergence between Dakkhini and Telugu. Karkhandari, the variety of Urdu spoken by the Karkhandars of Delhi, has been studied by Narang (1961) and later by Rauf (1997), whose work is an articulatory phonetic study of that dialect.

### 2.2. In Pakistan

In Pakistan too, there is a paucity of *linguistic* work on Urdu. Scholarship on Urdu, aside from literary studies, has been largely devoted to the extra-linguistic historical, political, and ideological issues associated with it. Masud Husain Khan's study is a summary of research on Urdu up until 1969, in which he says, "the main areas which have interested the Urdu scholars are lexicography, grammar-writing, and textual criticism. Descriptive analysis is of recent growth" (1969, 283). This summary prominently mentions the efforts of the Linguistic Research Group of Pakistan organized by Anwar S. Dil, which culminated in his publication of three books (1963, 1964, 1965). Khan notes in 1969 that there is no chair or institute of linguistics devoting itself to Urdu language studies; rather, any linguistic studies of Urdu are attached to Urdu Departments, usually headed by literary scholars. He says, "Under these circumstances they have difficulty in doing justice to either linguistics or literature" (283). Dil (1969) gives a chronological account of Pakistani personalities and institutions involved in linguistic studies in Pakistan up to 1969. Tariq Rahman's 1998 report contains a summary discussion of linguistic work done in Pakistan up to that point. In it he concludes depressingly, "Pakistan is perhaps the most backward country of South Asia in the field of linguistics" (192). Bashir (2006) treats linguistic work in Pakistan and on Pakistani languages subsequent to Rahman (1998). New developments, including the establishment of a Department of Pakistani Languages at Allama Iqbal Open University and the advances being made in computational linguistics, are discussed. This report stresses the importance of technology in the current advances, and concludes on a somewhat more optimistic note than Rahman did in 1998.

### 2.3 Outside of South Asia

Paradoxically, but not surprisingly given the issues discussed below in Section 3, most of the linguistic work on Urdu has been done by scholars based outside of South Asia.

Some representative studies of Urdu are listed here, in rough chronological order. Donald Becker, at the University of Wisconsin, worked on Urdu phonology (Becker and Narang 1971), developed a computer font for Urdu, and published a reverse dictionary of Urdu (1980), an early application of the newly emerging computer technology to lexicography. Becker's dictionary has proved an invaluable tool for this author. Azim (1978) is a Columbia University Ph.D. dissertation on the verb system of classical Urdu. Tuite, Agha, and Graczyk (1985), an early article on Urdu semantics and typology, focuses on form and function in verb conjugations. A few of the works of Anjum Saleemi, one of the few Pakistanis, and perhaps the first, to work in theoretical linguistics, deal with Urdu (1994a, 1994b, 2004). Most of his work is in the generative framework and has focused on questions of universal grammar and language acquisition (e.g., 1992). Miriam Butt, starting with a Ph.D. dissertation on complex predicates in Urdu (1995), has been working mainly in the Lexical Functional Grammar (LFG) syntactic framework since then, drawing her data from what she specifically calls "Urdu." Some of her earlier publications are Butt and King (1991), Butt and Geuder (2001, 2004), and Butt and Sadler (2003). More recently she has done computational work on Urdu (e.g., Bögel et al. 2007; Butt, King and Roth 2007), and collaborated with the Centre for Urdu Language Processing in Lahore on computational projects including a 2003 Summer School on Morphology and Syntax of Urdu for Computational Linguists and development of a machine translation system. Bashir has three contrastive studies of Urdu and other Pakistani languages (1991a, 1991b, 1991c) and two articles on recent language change in Urdu (1999, 2006). Hussain (1997) is a Northwestern University Ph.D. dissertation on the phonetic correlates of lexical stress. Hameed (2004) is a phonological study of Lucknow Urdu, and Ahmad (2007), a University of Michigan dissertation, is a quantitative sociolinguistic study of chronological changes in indexicality values of Urdu in Old Delhi.

### 3. Reasons for This Disconnect

A confluence of historical and cultural currents affecting linguistics on the one hand and Urdu on the other resulted in the gulf I am addressing in this paper. The cultural currents affected both the field of linguistics and of Urdu studies.

### 3.1 Cultural Developments

#### 3.1.1 *What happened in the field of linguistics?*

After the Chomskyian revolution, linguistics fell prey to a sense of restrictiveness, when syntax and formalism came to be valorized above other aspects of linguistic analysis. This state of affairs and the resulting acrimony within the field obtained in the U.S. from the 1960s until perhaps 1990, and was, unfortunately, responsible for turning some promising young scholars away from academic linguistics.<sup>5</sup> This, however, did not happen in Europe. Roman Jakobson, who was simultaneously Slavicist, philologist, phonologist, folklorist, and literary theorist, addressed the relationship between poetics and linguistics as follows: “Poetics deals with problems of verbal structure, just as the analysis of painting is concerned with pictorial structure. Since linguistics is the global science of verbal structure, poetics may be regarded as an integral part of linguistics” (1960, 350). And further:

Insistence on keeping poetics apart from linguistics is warranted only when the field of linguistics appears to be illicitly *restricted* [emphasis mine], for example, when the sentence is viewed by some linguists as the highest analyzable construction or when the scope of linguistics is confined to grammar alone or uniquely to nonsemantic questions of external form or to the inventory of denotative devices with no reference to free variations.

(*ibid.*, 352)

And also:

If there are some critics who still doubt the competence of linguistics to embrace the field of poetics, I privately believe that the poetic incompetence of some bigoted linguists has been mistaken for an inadequacy of the linguistic science itself. All of us here, however, definitely realize that a linguist deaf to the poetic function of a language and a literary scholar indifferent to linguistic problems and unconvertant with linguistic methods are equally flagrant anachronisms.

(*ibid.*, 377)

Jakobson also addresses the issue of excessive normativeness in literary studies:

Unfortunately the terminological confusion of “literary studies” with “criticism” tempts the student of literature to replace the description of the intrinsic

---

<sup>5</sup>Some aspects of the field-internal conflicts are sometimes called the “linguistics wars” (Harris 1993).

values of a literary work by a subjective, censorious verdict. The label “literary critic” applied to an investigator of literature is as erroneous as “grammatical (or lexical) critic” would be applied to a linguist. Syntactic and morphologic research cannot be supplanted by a normative grammar.

(*ibid.*, 352)

Examples of the happy marriage of poetics and linguistics are not hard to find: Jakobson himself; Watkins’ work (1995) using historical linguistic methods to reconstruct and analyze Indo-European poetry; the work of Paul Friedrich, who is both linguist (1970, 1986) and poet (2010). A recent article by Michael Wagner and Katherine McCurdy argues, based on experimental evidence, that:

The restrictions on identical rhymes across languages constitute further evidence that a better understanding of the linguistic system of a language can illuminate the study of poetry and vice-versa, as advocated by Jacobson, 1960, and [...] Kiparsky (1973).

(2010, 174)

### 3.1.2 *What happened in Urdu Studies?*

Several factors have contributed to the relative neglect of Urdu by linguists. Extremely important is the structural similarity of Urdu and Hindi: the syntax of Urdu and Hindi is almost identical. Since modern linguistics has until recently been dominated by syntax, and the syntax of Hindi and Urdu is so similar, it has been generally felt that linguistic studies of Hindi apply to Urdu as well. This is despite significant differences in phonology, morphology and lexis (cf. Khan 1989).<sup>6</sup> Most scholars working on Urdu or Hindi title their works using the term “Hindi-Urdu” to reflect this structural similarity. Among these, since the field of linguistics got off to a much earlier start in India than in Pakistan, the vast majority of linguistic studies are by scholars working with Hindi data or sources.

Some characteristics of Urdu literary culture have also played their part—importantly, a tendency to prescriptivism. Shamsur Rahman Faruqi says:

One manifestation of the new Urdu culture was its almost morbid obsession with “correctness” in language. Undue—and sometimes almost mindless—emphasis on “correct” or “standard, sanctioned” speech in poetry and prose, and even in everyday converse, is one of the most interesting and least understood aspects of Urdu culture from the mid-

---

<sup>6</sup>An extreme example of this can be seen by comparing two dictionaries of linguistics terminology: one for Hindi almost entirely based on Sanskrit roots (Nardella 2008) and one for Urdu (A‘vān 1995) based on Perso-Arabic roots.

eighteenth century onward.<sup>7</sup> Persian's immense prestige ("Persian" here includes Arabic) may account for a part of this emphasis.

[...]

[...] many of the taboos that originated in the early nineteenth century are still in place. In theory, and also to a large extent in practice, Urdu literary idiom remains the most restrictive kind imaginable.

(2001, 152, 155–56)

These remarks remind us of Jakobson's comment on excessive normativeness in literary studies. A recent comment from the Internet mailing list URDULIST (April 22, 2010), illustrates this attitude. Acknowledging that the Urdu script has been kept vital in Pakistan, the writer of the post opines that: "[...] the Urdu that is spoken in Pakistan is now heavily corrupted by local dialects..." Since most linguists abjure prescriptivism, this characteristic of Urdu literary culture has made it unattractive to some linguists.

Urdu is generally associated in most people's minds with literature, especially poetry. In this author's experience, wanting to be able to read and understand Urdu poetry is the second most frequent reason cited by prospective students for wanting to study Urdu. A "quick and dirty" Google search (December 22, 2010) retrieved 1,890,000 hits for "Urdu poetry" and 2,630 for "Urdu linguistics." According to Salman Khurshid, Urdu is "stereotyped as the mellifluous language of art and literature while Muslims are caricatured [*sic*] in Bombay films. [...] So, in free India, Urdu has never been recognized as a functional language" (2006, ix). It is possible that, internalizing this perception, linguists, who mostly need prose texts or natural oral discourse for their work may avoid Urdu. An additional, more recent, cultural factor is the declining social prestige of Urdu in India. According to Rizwan Ahmad, the status of Urdu in India has changed, at least in Delhi.

To the first generation Muslims and Hindus born *before* the Partition of India in 1947, Urdu indexes education and cultural refinement. To the second generation born *after* 1947, Urdu indexes an exclusive Muslim identity. The indexicality of Urdu, however, undergoes a reconfiguration again among the third generation Muslims who were born after the early 1980's; to them, Urdu indexes a poor, uneducated, and conservative Muslim identity. [...]

[...]

Among Old Delhi youth, Urdu indexes backwardness and lack of

---

<sup>7</sup>The writer of the current paper finds this remark immensely intriguing and hopes that scholars of Urdu take it up as a question for research.

education—a language that does not going along with [*sic*] the modern identity of the youth.

(2007, xxiii, 105)

### 3.2 Political-historical factors

Paradoxically, Urdu has been regarded as a language of outsiders both in India and in Pakistan.

#### 3.2.1 *In India*

First, let us consider India. Barbara Metcalf quotes the following line from an unpublished poem of Rāshīd Banārsī: *Agar Urdū pe bhī ilzām hai bāhar se ānē kā, to phir Hindūstān kis kā vaṭan hai, ham nabīñ samjḥē* (If Urdu is accused of being an outsider, then I don't understand, whose homeland is India?) (2006, 63). This characterization of Urdu as a “language of outsiders” is related to the widely held idea that Urdu is (exclusively) the language of Muslims. Lelyveld notes: “In post-independence India, Urdu is associated with Muslim identity,” and further, “It [Urdu] is not only almost universally identified as a language for a regionally unspecified Muslim population, it is also the would-be, official, non-regional language of a foreign country, Pakistan” (1993, 682). Jinnah's insistence that Urdu was the language of Muslims and that Urdu and only Urdu be the national language of Pakistan reinforced this development. Ahmad shows that the association of Urdu with Muslims was strengthened after 1947 and the adoption of Urdu as the national language of Pakistan. As noted earlier, according to him, the generation of people born and raised before 1947 did not make this exclusive association of Urdu with Muslims (2007, xxiii).

#### 3.2.2 *In Pakistan*

In Pakistan, Urdu was initially and for a long time seen by the majority of the people as a language of outsiders, the Urdu-speaking *muhājirs* (refugees, immigrants) who were perceived as (actually) Indians. It was the language identified with the ethnic group of North-Indian Muslims who dominated Pakistani political life for the first two decades after 1947.

Ayres points out that

[...] Pakistan's emergence into the world of nation-states took place against an assumption of cultural and linguistic national consciousness which located this new nation's historical narrative primarily in lands which remained in India. This historical narrative privileged the literary traditions of the Urdu language as the exemplar and indeed repository of Muslim consciousness.

(2009, 188–89)

The resultant privileging of Urdu by the state led to violent opposition to it by various ethnic groups, especially in Bengal and Sindh.

Linguistics in Pakistan is at a rudimentary stage of development compared to India. There are several reasons for this. From Pakistan's inception, the "one nation, one language" model was accepted. Ayres quotes Jinnah's insistence on Urdu as the national language from a 1948 speech on "National Consolidation" in Dhaka:

[...] But let me make it very clear to you that the State Language of Pakistan is going to be Urdu and no other language. Any one who tries to mislead you is really the enemy of Pakistan. Without one State Language, no Nation can remain tied up solidly together and function. Look at the history of other countries. Therefore, so far as the State Language is concerned, Pakistan's language shall be Urdu.

(Qtd. in *ibid.*, 42)

Thus, at a time when consolidation of a national identity with Urdu as its government-supported medium of expression was preoccupying the government and many intellectuals, the study of linguistics was initially perceived as potentially encouraging centrifugal tendencies.

Now, however, with the passage of over sixty years, Urdu has gained considerable traction in Pakistan. It is used almost universally in one form or another as a link language, and has achieved the status of a successfully functioning lingua franca and national language. Since it is the medium of education in the entire country except Sindh, school-going children do learn it. As Urdu gains currency, there appears to be a softening of attitudes and a pulling back from the "one nation, one language" idea. An article by Jonaid Iqbal in the *Dawn* newspaper on 26 May 2010 reported that at a seminar on "The Role of Pakistani Languages in Nation Building,"<sup>8</sup> National Language Authority Chairman Iftikhar Arif said that the Father of the Nation was quoted out of context when he said no nation could achieve unity without a national language. Iftikhar Arif said the Quaid, in his speech during his visit to Dhaka in March 1948 emphasizing Urdu as the national language, had said the provinces had the full right to determine the language of their provinces, but people tend to forget this prominent part of the Quaid's speech.<sup>9</sup> Even more

---

<sup>8</sup>Note the wording of the seminar theme: "Pakistani languages" [plural] rather than "national language." The *Dawn* article also reports that "they [members of the intelligentsia] also asked the government to take practical steps for protection and development of 39 major languages of the country."

<sup>9</sup>The further context referred to by Iftikhar Arif begins:

recently, *Dawn* reported on 26 January 2011 that a private bill, Constitution (Amendment) Bill 2011 had been introduced in the National Assembly proposing to replace Article 251 of the Constitution, which specifies only Urdu as the national language, with a new article designating eight languages (Balochi, Panjabi, Pushto, Shina/Balti, Sindhi, Saraiki and Urdu) as national languages. Regardless of whether this bill ultimately is adopted or not, the very fact that it has been proposed marks a major shift.

#### 4. Recent Linguistic Scholarship on Urdu in Pakistan

Since 1947, the availability of training in linguistics has been severely lacking in Pakistan. Until 2009 there had been no department of linguistics in Pakistan.<sup>10</sup> Linguistic work on Urdu in Pakistan has been mostly moti-

---

Whether Bengali shall be the official language of this Province is for the elected representatives of the people of this Province to decide. I have no doubt this question shall be decided solely in accordance with the wishes of the inhabitants of this Province at the appropriate time.

Let me tell you in the clearest language that there is no truth that your normal life is going to be touched or disturbed so far as your Bengali language is concerned. But ultimately it is for you, the people of this Province, to decide what shall be the language of your Province. But let me make it very clear to you [...].

(Jinnah 1989, 183)

<sup>10</sup>J.R. Firth showed an early interest in Urdu and linguistics in India and Pakistan. In a 1957 address to the Philological Society on "Applications of General Linguistics," describing a visit to India and Pakistan, he said:

India has taken up American linguistics, including large doses of phonemics, in the hope of carrying out new linguistic surveys with a view to the enrichment of the national language from dialect sources, and to promote some convergence at any rate in vocabulary of the principal languages. They have worked on technical terminology in the sciences and on nomenclature and phraseology for the administrative and defence services. My impression is that they still have a long way to go even in the preliminary exploratory and learning period. And phonemics, like patriotism, is not enough. In Pakistan, the language problems are not so vast or so intricate, and they are not yet committed to any extensive programme of linguistic research. But at a recent Conference in Karachi which I attended, along with American representatives, a special Committee composed of two Vice-Chancellors, two senior educationists and three Pakistani linguists, decided to recommend the gradual establishment of at least two University departments of general linguistics, and strongly urged the training of suitable young scholars abroad.

(In Palmer 1968, 133)

vated and shaped by two factors: that Urdu is the national language and link language for the country, and that there is a desire to improve English-language teaching. Classes in applied linguistics have most often been taught in the contexts of English as a Second Language (ESL) or sociolinguistics, usually as elective courses in departments of English. In the absence of academic training in general or theoretical linguistics, scholars who wanted to study and write about Urdu were, in the opinion of Tariq Rahman, perforce limited to comparative or sociolinguistic research. The prolific work of Tariq Rahman himself on issues of language and power, politics and education (see e.g., 2002, 2010) constitutes a major component of this work. Numerous works on establishing the pedigree of Urdu are cited by Rahman, who considers that, “One major theme of people writing in [this] tradition is discovering the origin, the family, and the roots of a language. In the case of Urdu, this is an obsession” (1998, 185). Many such works are ideologically and politically motivated; they are too numerous to list.

#### 4.1 Encouraging Signs

Recently in Pakistan there is a new recognition of the lack of linguistic scholarship in the country. A 2009 article by Rauf Parekh in *Dawn* honoring Abu'l-Laiš Şiddiqī laments this fact. Parekh points out Pakistan’s backwardness in this field compared to India and criticizes “certain scholars,” who, preferring literary criticism, talk about linguistics with “scorn and derision.” Parekh’s perception is very similar to the feeling that has prompted me to write this paper. A recent *Jañg* newspaper article by Kishvar Nāhīd (2010) leads me to think that perhaps this recognition is becoming more widespread. With the advent and increasing availability of the Internet, research materials are more available to students and researchers. This has led to a growing interest in linguistics among students. I often see email questions from students to international mailing lists asking for guidance on linguistics-related thesis topics.

Until recently there had been no department of linguistics in any Pakistani university. However, a recent initiative by Dr. Navīd-e Rāḥat, Dean of the Faculty of Social Sciences at Quaid-i-Azam University, Islamabad, to establish a Department of Linguistics is highly encouraging. Classes began in February 2010 in the M.Sc. Linguistics program. Chances of its success are brighter now than they have been for previous such attempts, and I look forward to learning how this effort progresses.

#### 4.2 Computational Linguistics

For the past nine years or so, a new wave of studies in Urdu linguistics

has been gaining momentum. It is important and historically interesting that the *information technology* sector, specifically computer science, rather than any humanities-related department, is leading the way in these new developments. This sharp turn in the trajectory of linguistic studies of Urdu has been brought about by the worldwide information technology revolution and the keenly felt needs to make Urdu usable as a language of electronic communication, and to localize essential forms of software. This is another of the perhaps unanticipated consequences of establishing Urdu as the state-sponsored language of a nation-state. Hussain (1997), Lodhi (2004), and Rizvi (2007), for example, are three important recent Ph.D. dissertations on Urdu. Hussain is an instrumental investigation of the effects of stress on the phonetic properties of Urdu vowels and consonants, and the theoretical implications of these phonetic changes. Lodhi is a computationally oriented work aimed at developing an Urdu character pattern recognition system which can classify patterns even under non-optimal conditions. Rizvi “proposes an algorithm for parsing Urdu sentences based on closed-word-classes” (2007, vi). Humayoun, Hammarström and Ranta (2007) presents work on software for Urdu grammar.

Several institutional initiatives in Pakistan are directed toward these objectives. The National Language Authority (Muqtadira Qaumī Zabān) in Islamabad has published a series of dictionaries of various professional terminologies (e.g., Sabzwārī 1995) and several bi- or multi-lingual Urdu-Language X dictionaries (e.g., Khaṭṭak 1987). It is working on a Nasta‘īq font for Urdu and other Pakistani languages and on machine translation projects. It has plans to create a corpus of Urdu, including both contemporary and historical texts, referred to as the “Urdu Database” and “Urdu Databank.”

The Centre for Research on Urdu Language Processing (CRULP) was established in July 2001 at the National University of Computer and Emerging Sciences in Lahore to conduct research and development in three areas, including speech processing, computational linguistics and script processing. Their speech processing lab worked on developing an Urdu speech interface for computers, which, crucially for the field of linguistics, involves basic research in phonetics, phonology, and speech systems, as well as other more practically oriented topics. The language processing and computational linguistics sections worked on producing language applications for Urdu, including a machine translation system, Urdu grammar and spell checkers, and Urdu lexicon development. An online Urdu dictionary has been produced, which in addition to word meanings includes essential grammatical information, and historical examples of usage. Recently, Urdu localization of the SeaMonkey Internet

application suite, and work on a Microsoft Vista Urdu Language Interface Pack was announced. Their most recent project is the Dareecha project, which aims to provide Information and Communication Technology (ICT) access and local language content generation capability in rural schools. The goal is for school children to be able to start with accessing the Internet at the beginning of the training program, and eventually reach a level by the end of the program where they are producing their own content and making it available on the Internet. All training in the Dareecha program will be conducted for Urdu-localized software. The CRULP research group has recently moved to the University of Engineering and Technology (UET) in Lahore, as the Centre for Language Engineering, Al-Khwarizmi Institute of Computer Science, under the leadership of Sarmad Hussain.

The Computer Science Department at Peshawar University has produced work on computational problems in both Urdu and Pashto. In 2008 a Society for Natural Language Processing was established, which is now headquartered at the Centre for Language Engineering, UET, Lahore. Its main objectives are to coordinate the multiple efforts at computational linguistics going on in Pakistan.

Work on Optical Character Recognition (OCR) software for Nasta'liq script is under way in several places in Pakistan including: the University of Engineering and Technology in Lahore; the Computer Science Department at NED [Nadirshaw Edulji Dinshaw] University of Engineering & Technology in Karachi (Sattar et al. 2008); the Ghulam Ishaq Khan (GIK) Institute of Engineering Sciences and Technology in Topi, Khyber Pakhtunkhwa (Husain and Amin, 2002); and the College of Signals, National University of Science & Technology (NUST), Rawalpindi (Rehman 2010). Working at SUNY Buffalo in the U.S., Mukhtar, Setlur and Govindaraju (2009) discuss and include a bibliography of previous work on Nasta'liq OCR, mentioning in particular the need for a corpus of handwritten Urdu.

### 5. New Developments in Urdu in Pakistan

The transplantation of Urdu from its native soil in India to that of Pakistan has given rise to rapid linguistic change and multiple new situations that are fertile ground for research, particularly in the areas of language contact and historical linguistics. Some scholars have expressed a tinge of regret at the passing of an old order. For example, Metcalf quotes M. U. Memon from a 2001 article in *The News* as saying:

No matter how one tries to squash the initiatives to promote it, [Urdu] will

grow if only because our [Pakistani] national life requires it. So there is no danger of Urdu disappearing. But what kind of Urdu would that be? Not the language that will make the subtleties of the poetry of a Mir and Ghalib accessible or result in a renaissance of Urdu literature. It will be a functional language.

(2006, 66)

On the other hand, Intizār Ḥusain, writing in *Akbbār-e Urdū* in 1998, in an article later republished in the *Annual of Urdu Studies* (2000) and further cited in Yaqin (2006), finds cause for optimism. In Yaqin's words, he feels that, "the cultural tradition of Urdu lies in its shifting regional locations. According to him, this language cannot be associated with one region and one culture because it is by nature hybrid and adaptable to new regions" (119). After this general statement, Ḥusain goes on to discuss the inevitability of language change and mixing, and points out the fallacy of "linguistic purity." A language, if it is to remain living, cannot be bound by nostalgic memories of a premodern past, agrees Yaqin.

### 5.1 Social Effects of Lingua Franca Usage

One potential area of research that immediately suggests itself concerns the effects of the development and use of a lingua franca—both on society and on the languages concerned. The following general observations are relevant to these questions, particularly with reference to the lingua franca function of Urdu in Pakistan. It is possible that a diglossic situation may develop, in which one language is used for "high" functions (i.e., in formal interactions such as official discourse, schools, and written communication), and the other language is used for so-called "low" functions (informal, domestic, quotidian, mainly oral interactions). As a result the high language is often held to be more logical and aesthetic, and associated with literary heritage; it is the one that is acquired through formal education. The resulting use of a "high" variety as the medium of instruction, especially if it is used as the sole medium throughout a child's educational career, may have severe consequences for pupils' social, cognitive, and academic development. One need only read the studies on the value of mother-tongue education to see the implications of this. For example, Prew (2011) presents data on the relative success of South African schools where education is carried out in the mother tongue in the initial stages.

Another consequence is language shift, the situation in which people give up their original mother tongues in favor of the lingua franca, especially if it is a higher-prestige language. This can result in the loss of linguistic diversity. Mufwene, discussing globalization, language shift and

language loss, argues:

Thus, while English has shaken the position of French and other colonial European languages in these former exploitation colonies, it is far from endangering the indigenous ones. English is not really the “killer language” that non-global approaches to language endangerment have painted it to be, certainly not in relation to the indigenous languages of former exploitation colonies.

(2008, 251)

And further,

If anything endangers the ethnic vernaculars, it is not the “global” languages used at the top echelon of the multinational companies, rather it is the urban vernaculars and regional lingua francas [...] that do.

(*ibid.*, 255–56)

## 5.2 Language Change

Urdu’s position as national language and use as lingua franca in Pakistan has resulted in rapid and massive changes, both in Urdu and in the other languages with which it is in contact.

### 5.2.1 *Changes in Urdu in Pakistan*

It is well known that the structures of local vernacular languages influence speakers’ productions in a lingua franca. Phenomena of this kind are often discussed under the rubrics of “transfer” or “substratum effects.” A classic example of these phenomena is the case of Urdu in Kupwar (India), which has changed under the influence of Kannada and Marathi in features discussed in Gumperz and Wilson (1971). Urdu in Pakistan is changing rapidly, in grammar, phonology, and lexis, to the extent that many people now consider Pakistani Urdu as a new variety. This was noted as early as 1966 by Azhar, who observed differences between Indian and Pakistani Urdu and argued that absorption of words from the indigenous languages of Pakistan was a natural process and would, sooner or later, result in all Pakistanis taking ownership of Urdu. A set of articles written in connection with the fiftieth anniversary of Pakistan’s creation (Durrani 1997) discusses these same points, emphasizing the revitalizing aspects of language change. Given the existence of new diasporic varieties, perhaps Urdu itself will need to be studied as a pluricentric language, in addition to its participation in the Hindi-Urdu continuum. A few of these changes have been noted and studied, but mostly these phenomena remain as yet unquantified and unanalyzed.

One example of such change is the reanalysis of the ergative marker

*ne* to indicate agency and even future intention. An instance of this change is shown here as (1).<sup>11</sup> In this example, an exchange between speakers A and B, in which A is trying to persuade B to go for coffee with him. The reply *maiñ nē nabīñ jāñā* (“I won’t go”) is used as a forceful indication of speaker B’s (agentive) unwillingness to go.

(1) A: *maiñ tumbēñ kāfi pilātā - baṛī zabardast -*  
 I you(DAT) coffee drink(CS) very wonderful-  
*čalō uṭhō*  
 come.on get.up  
 ‘I’ll take you for coffee—really great (coffee), come on, get up.’

B: *maiñ nabīñ jāñgī*  
 I not will.go  
 ‘I won’t go.’

A: *dō ghanṛē mēñ tumbārā daftar bhāg tō nabīñ*  
 two hours in your office run.away TOP not  
*jāēgā*  
 will.go  
 ‘Your office isn’t going to run away in a couple of hours.’

B: *maiñ- nē nabīñ jāñā*  
 I=ERG not go(INF)  
 ‘[I told you] I don’t want to/won’t go.’

(Bashir 1999, 17;  
 qtd. from Pakistan Television [PTV] drama  
*Tanbā’iyāñ* (Loneliness), by Ḥasīna Muṣīn)

This change has been long noted and much discussed in the literature<sup>12</sup> and has by now become emblematic of Pakistani Urdu as a separate variety.

A second noticeable change in progress is that the category of grammatical gender is weakening. There is increasing uncertainty about gender assignment resulting from differences between inherited grammatical gender patterns in Urdu and the languages with which it is interacting in Pakistan.<sup>13</sup> Some of these languages do not have grammatical gender

<sup>11</sup>Abbreviations for linguistic terms used here are: CS-causative; DAT-dative; ERG-ergative; FUT-future; INF-infinitive; TOP-topic marker. Retroflex consonants are indicated with an underdot; long vowels with a macron above and nasalized vowels by a ñ following the vowel, according to the *AUS* style sheet. As far as possible, these representations are transcriptions, not transliterations.

<sup>12</sup>See Bashir (1999, 12–15) regarding discussions of this feature.

<sup>13</sup>S.A. and M.A. Khan (2009) describes a computational approach to the problem of gender assignment in Urdu.

(e.g., Balochi or Khowar), and in some, gender patterns operate differently than they do in Urdu (e.g., Pashto). Bashir discusses and illustrates how gender assignment patterns differ in Urdu and Pashto (1991a, 23–33, 247–55). This results in a tendency to assign many nouns a default masculine singular gender, and may lead to the eventual loss of the category. This is what happened historically in Iranian languages such as Persian, and in other Indo-Aryan languages such as Khowar or Bengali. Changes in gender assignment of some words over the last 115 years can be observed by comparing lexical entries in an older dictionary, such as Platts' dictionary from 1884, with entries in the newer *Kitabistan's 20th-Century Standard Dictionary, Urdu into English* (Qureshi 1971). For example, *fikr* (thought, worry) is feminine in Platts' but masculine in Qureshi; similarly, *qaum* (people, nation) is given as feminine or masculine in Platts', but as masculine in the *20th-Century Standard Dictionary*.

Phonological changes are also noticeable, some quite general and some associated with different regions. Voiced aspirates are losing aspiration across the spectrum of Urdu speakers; for example, it is common to hear the word *Bhārāt* (India) pronounced as “Bahārat,” with a sequence of /b/, short vowel, consonantal /h/, replacing the voiced aspirate /bh/. Other, less generalized changes include the low tone developing on long vowels (especially ā) in the Urdu of many Panjabi speakers, or the unique sentence intonation patterns of Urdu speakers whose first language is Pashto. It may be that, rather than a singular Pakistani Urdu, multiple Pakistani Urdu(s) are being born.

### 5.2.2 Changes in Other Languages

The influences of Urdu on local languages in Pakistan are also numerous and pervasive. These contact-induced effects can be studied from the points of view of traditional historical linguistics, or areal linguistics. Many of these changes are lexical, but some involve changes in the grammar, for example the increasing use of specifically progressive forms in Balochi (Bashir 2008, 77–78) and Brahui (Bashir 2010, 1). Bashir (2007) is a study on contact-induced changes in Khowar lexicon, phonology and syntax. The phenomena of code switching and code mixing are familiar mostly from discussion of the Urdu-English situation, but they also occur between Urdu and Panjabi, for example, and seem to be ubiquitous in South Asia. The Urdu-English case is fairly well studied, but one wonders how this phenomenon is manifested in a multilingual situation like Quetta. These facts of language change can lead to a rich development of dialect studies, both within Pakistan and in various diasporic communities. Similarly, the divergence of Pakistani Urdu(s) from Indian

Urdu, and of Urdu from Hindi<sup>14</sup> are rich areas for study. Local varieties of Urdu spoken in various diasporic communities can shed light on how speakers construct multiple identities in an increasingly interconnected world.

### 6. Desiderata and New Directions for Study

The “information technology revolution” has given new life to the study of linguistics in Pakistan. However, the people driving this new research are generally more computer scientists than linguists. Hence, there is still a gulf, albeit of a new kind, between computationally-oriented Urdu studies and other fields within linguistics. What is now needed is to develop a cadre of general linguists in Pakistan whose skills are at the same level of sophistication as those of the computer scientists. After that, the next step, it seems to me, will be to begin to develop those aspects of linguistics which can bridge the gap between the humanities, including both linguistics and literary studies, and the sciences. Several areas of investigation suggest themselves.

Important sociolinguistic questions beg for attention. What are the social effects of the use of a dominant, non-native language in different communities? To what extent is Urdu becoming “native”? To what extent is language shift taking place in different linguistic communities? The situation with Panjabi is well known, where in Punjab, language shift is taking place at an increasing rate in the urban centers. Urdu is being increasingly adopted, largely as a result of Panjabi-speaking families speaking Urdu with their children to give them a “head start” in school. This does not happen with Pashto speakers, for several reasons, including linguistic structural factors. This is an area that demands study, since accelerating language shift is leading to a reduction of linguistic diversity within the country.

Corpus development is an area which both demands computational skills and tools and can be applied to the kinds of questions scholars of literature often ask. For instance: “What are the characteristics (lexical, grammatical, rhetorical) of writer X’s style?” “How has the use of construction Y varied over time?” “How has the meaning of word Z changed over time?” Efforts in corpus building have been made for Urdu—the EMILLE Project in England (McEnery et al. 2000; Hardie 2003; 2005); in India (Dash 2004); in Lahore (Ijaz and Husain 2007) and in Islamabad with

---

<sup>14</sup>Iqtidar Khan (1999) is a study of Hindi and Urdu differences at that time. A study on subsequent developments would add historical depth to these questions.

the National Language Authority's Urdu Database project. The existence of good corpora—both oral and written—will open the doors to all kinds of corpus-based research, with applications in several fields like literary study, historical linguistics, and education. A glance at the journal *Literary and Linguistic Computing* (e.g., Baker et al. 2004) will give an idea of the possibilities.

Discourse analysis is another area in which literary scholars have a deep interest. This is the study of language as it is used to communicate between people. It usually involves structural units larger than the sentence, and can include topics like textual cohesiveness (e.g., anaphora), rhetorical strategies, information structure, and figurative language. Existing research in these areas ranges from studies of pronominal anaphora (Khan, M.A. et al. 2007) to research on prosodic structure (e.g., Damron 2004) to studies of comparative Urdu and English rhetorical structure on the lines of Kachru (1988), a study of comparative rhetoric for Hindi and English. Akram (2008) is a comparative study of speech acts in Urdu and English.

A plethora of second language acquisition studies involve Urdu as the L1 and some other language, usually English, as the L2. Those on Urdu as the L2 are far fewer. Bashir (1991a, 1991b, and 1991c) are pedagogically oriented studies directed to teachers of Urdu as a second language in Pakistan.

These changes in the function, status and relationships of Urdu—in India, in Pakistan, and in other diasporic communities—force us to re-examine the notions of “first language,” “mother tongue” and “native speaker.” Recent interrogation of the notions of “mother tongue” and “native speaker” has grown largely out of the study of “World Englishes,” often in the context of ESL teaching (e.g., Rampton 1990), but the result of this new thinking is highly relevant to theorizing the position of languages such as Hindi in India and Urdu in Pakistan.<sup>15</sup> With regard to the Pakistani situation, are those children who grow up in households where the parents speak Panjabi among themselves but Urdu to the children, and the children speak Urdu in school and among themselves but (sometimes) Panjabi with their parents, native speakers of Urdu? What is their mother tongue? Their first language? Their second language? Given that the notion of “native speaker” carries overtones of ownership over a language and authority over other speakers as arbiters of usage,<sup>16</sup> these

---

<sup>15</sup>LaDousa (2010) is a discussion of the concept of “mother tongue” in the context of the schools in Banaras.

<sup>16</sup>Mufwene (1998) argues that proficient speakers, whether “native” or not, are

questions have a particular urgency in a state where issues of language and power are so historically fraught, and where the essentializing of identities is a persistent tendency. As Annamalai succinctly noted, “In a multilingual country nativity of language is a variable construct shifting with political perceptions” (1998, 154). The linguistic situation in Pakistan is in such rapid flux that it is difficult to decide the answers to such questions, and one wonders whether and to what extent such notions are even relevant for India and Pakistan? Ansaldo argues that in

[M]ultilingual contexts in which different languages are negotiated on a daily basis, and where language contact and contact languages are ubiquitous [...] the alignment between language and identity is complex, continuously shifting and not easily captured in terms of mother tongue or nativeness. In this sense, multilingual ecologies question the notion of mother tongue and its implicit and explicit role in our current theories of language.

(2010, 615)

Overall, the emerging consensus seems to be that a simplistic binary distinction between native speakers and non-native speakers needs to be replaced with multiple and “fuzzier” categories which reflect the complex and changing realities of multilingual societies.

It is possible that the complex new realities of Urdu will pique the interest of linguists, and the analytical tools linguistics can offer will attract scholars of Urdu literature. □

### Works Cited

- Ahmad, Rizwan. 2007. *Shifting Dunes: Changing Meanings of Urdu in India*. Ph.D. diss., University of Michigan.
- Akram, Muhammad. 2008. “Speech Acts: A Contrastive Study of Speech Acts in Urdu and English.” *Asian EFL Journal* 10(4, Article 8) [[http://www.asian-efl-journal.com/December\\_08\\_ma.php](http://www.asian-efl-journal.com/December_08_ma.php)].
- Annamalai, E. 1998. “Nativity of Language.” In *The Native Speaker: Multilingual Perspectives*. Ed. by Rajendra Singh. New Delhi; London: Sage Publications.
- Ansaldo, Umberto. 2010. “Identity Alignment and Language Creation in Multilingual Communities.” *Language Sciences* 32(6):615–23.
- Arora, Harbir. 1986. *Some Aspects of Dakkhini Hindi-Urdu Syntax with Special Reference to Convergence*. Ph.D. diss., Delhi University.

---

the most appropriate arbiters of language usage.

- and K.V. Subbarao. 1989. "Convergence and Syntactic Reanalysis: The Case of *so* in Dakkhini." *Studies in the Linguistic Sciences: Papers In General Linguistics* 19(1):1–18.
- and K.V. Subbarao. 1988. "On Extreme Convergence: The Case of Dakkhini Hindi-Urdu." *Indian Linguistics* 49(1–4):92–108.
- Aʿvān, Ilāhī Baksh Akhtar. 1995. *Kashshāf Iṣṭilāḥāt-e Lisāniyāt* [Explanation of Linguistics Terms]. Islamabad: Muqtadira Qaumī Zabān [National Language Authority].
- Ayres, Alyssa. 2009. *Speaking Like a State: Language and Nationalism in Pakistan*. Cambridge; New York: Cambridge University Press.
- Aẓhar, A.D. 1966. "Pakistānī Urdū aur Hindustānī Urdū." *Nuṣrat* (Dec.):9–24.
- Azim, Abdul. 1978. *Systems of the Verb in Classical Urdu*. Ph.D. diss., Columbia University.
- Bailey, T. Grahame. 1939. "The Date of the Word 'Urdu'." *Journal of the Royal Asiatic Society of Great Britain and Ireland* 2:264–65.
- . 1934. "The Four Classes of Urdu Verbs." *Bulletin of the School of Oriental Studies* 7(3):503–7.
- . 1931a. "Urdu Grammatical Notes, I." *Journal of the Royal Asiatic Society of Great Britain and Ireland* 1:147–50.
- . 1931b. "Urdu Grammatical Notes, II." *Journal of the Royal Asiatic Society of Great Britain and Ireland* 2:418–20.
- . 1929. "The Meaning and Usage of Causal Verbs in Urdu and Panjabi." *Bulletin of the School of Oriental Studies* 5(3):519–22.
- . 1922. "The Pronunciation of Urdu and Hindi." *Bulletin of the School of Oriental Studies* 2(3):539–47.
- Baker, Paul [et al.] 2004. "Corpus Linguistics and South Asian Languages: Corpus Creation and Tool Development." *Literary and Linguistic Computing* 19(4): 509–24.
- Bashir, Elena. 2010. "Innovations in the Brahui Verb System." *Journal of South Asian Linguistics* 3(1):23–44.
- . 2008. "Some Transitional Features of Eastern Balochi: An Areal and Diachronic Perspective." In *The Baloch and Others: Linguistic, Historical and Socio-Political Perspectives on Pluralism in Balochistan*. Ed. by C. Jahani, A. Korn, and P. Titus. Wiesbaden: Reichert.
- . 2007. "Contact-Induced Change in Khovar." In *New Perspectives on Pakistan: Contexts, Realities and Visions of the Future*. Ed. by Shafqat Saeed. Karachi: Oxford University Press.
- . 2006. "Pakistan: Research and Developments in Linguistics and Language Study." In *The Yearbook of South Asian Languages and Linguistics 2006*. Ed. by Rajendra Singh. Berlin: Mouton de Gruyter.
- . 1999. "The Urdu Postposition *ne*: Its Changing Role in the Grammar." In *The Yearbook of South Asian Languages and Linguistics 1999*. Ed. by Rajendra Singh. New Delhi; London: Sage Publications.
- . 1991a. *A Contrastive Analysis of Balochi and Urdu*. Washington, D.C.; Peshawar, Pakistan: Academy for Educational Development.

- . 1991b. *A Contrastive Analysis of Brahui and Urdu*. Washington, D.C.; Peshawar, Pakistan: Academy for Educational Development.
- . 1991c. *A Contrastive Analysis of Pashto and Urdu*. Washington, D.C.; Peshawar, Pakistan: Academy for Educational Development.
- Becker, Donald A. 1980. *A Reverse Dictionary of Urdu*. New Delhi: Manohar.
- and Gopi Chand Narang. 1971. "Aspiration and Nasalization in the Generative Phonology of Hindi-Urdu." *Language* 47(3):646–67.
- Bögel, Tina [et al.] 2007. "Developing a Finite-State Morphological Analyzer for Urdu and Hindi." In *Finite-State Methods and Natural Language Processing*. Sixth International Workshop FSMNLP Potsdam, Germany. September 14–16. Revised papers. Ed. by Thomas Hanneforth and Kay-Michael Würzner. Potsdam: Universitätsverlag Potsdam, 2008.
- Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu*. Dissertations in Linguistics Series. Stanford, CA: Center for the Study of Language and Information.
- and Louisa Sadler. 2003. "Verbal Morphology and Agreement in Urdu." In *Syntactic Structures and Morphological Information*. Ed. by U. Junghans and L. Szucsich. Berlin: Mouton de Gruyter.
- and Tracy Holloway King. 1991. "Semantic Case in Urdu." In *Papers from the 27th Regional Meeting of the Chicago Linguistic Society (CLS)*. Ed. by L. Dobrin, L. Nichols, and R. Rodriguez. Chicago: Chicago Linguistic Society.
- and Wilhelm Geuder. 2004. "Light Verbs in Urdu and Grammaticalization." In *Words in Time: Diachronic Semantics from Different Points of View*. Ed. by R. Eckardt, K. von Heusinger, and C. Schwarze. Berlin: Mouton de Gruyter.
- and Wilhelm Geuder. 2001. "On the (Semi)lexical Status of Light Verbs." In *Semi-lexical Categories: The Function of Content Words and the Content of Function Words*. Ed. by Norbert Corver and Henk van Riemsdijk. Berlin; New York: Mouton de Gruyter.
- , Tracy Holloway King, and Sebastian Roth. 2007. "Urdu Correlatives: Theoretical and Implementational Issues". In *Proceedings of the LFG07 Conference*. Ed. by Miriam Butt and Tracy Holloway King. [http://csli-publications.stanford.edu/LFG/12/lfg07-toc.html].
- Damron, Rebecca. 2004. In *Discourse Across Language and Cultures*. Ed. by Carol L. Moder and Aida Martinovic-Zic. Amsterdam: Benjamins.
- Dash, Niladri S. 2004. *Language Corpora: Present Indian Need*. [http://www.elda.org/en/proj/scalla/SCALLA2004/dash.pdf].
- Dawn*. 2011. "NA Bill Proposes Eight 'National Languages'." January 26. [http://www.dawn.com/2011/01/26/na-bill-proposes-eight-national-languages.html].
- Dil, Anwar S. 1969. "Linguistic Studies in Pakistan." In *Current Trends in Linguistics, Vol. 5, South Asia*. Ed. by Thomas A. Sebeok. The Hague: Mouton.
- . 1965. *Studies in Pakistani Linguistics*. Lahore: Linguistic Research Group of Pakistan.
- . 1964. *Pakistani Linguistics (1963)*. Lahore: Linguistic Research Group of Pakistan.

- . 1963. *Pakistani Linguistics (1962)*. Lahore: Linguistic Research Group of Pakistan.
- Durrānī, ‘Aṭash. 1997. *Pākistānī Urdū kē Khad-o-Khāl* [Characteristics of Pakistani Urdu]. Islamabad: Muqtadira Qaumī Zabān [National Language Authority].
- Faruqī, Shamsur Rahman. 2001. *Early Urdu Literary Culture and History*. New Delhi: Oxford University Press.
- Friedrich, Paul. 2010. *A Goldfinch Instant: Concord to India Haikus*. Chicago: Virtual Artists Collective.
- . 1986. *The Language Parallax: Linguistics, Relativism and Poetic Creativity*. Austin: University of Texas Press.
- . 1970. *Proto-Indo-European Trees*. Chicago: University of Chicago Press.
- Gumperz, John J. and R. Wilson. 1971. “Convergence and Creolization: A Case from the Indo-Aryan/Dravidian Border in India.” In *Pidginization and Creolization of Languages*. Ed. by Dell Hymes. London: Cambridge University Press.
- Hameed, Shabana. 2004. “Interaction of Physiology and Communication in the Makeup and Distribution of Stops in Lucknow Urdu.” In *Cognitive and Communicative Approaches to Linguistic Analysis*. Ed. by Ellen Contini-Morava, Robert S. Kirsner, and Betsy Rodríguez-Bachiller. Amsterdam; Philadelphia: Benjamins.
- Hardie, Andrew. 2005. “Automated Part-of-Speech Analysis of Urdu: Conceptual and Technical Issues.” In *Contemporary Issues in Nepalese Linguistics*. Ed. by Y. Yadava [et al.]. Kathmandu: Linguistic Society of Nepal.
- . 2004. The Computational Analysis of Morphosyntactic Categories in Urdu. Ph.D. diss., University of Lancaster.
- . 2003. “Developing a Tag-Set for Automated Part-of-Speech Tagging in Urdu.” In *Proceedings of the Corpus Linguistics 2003 Conference*. Ed. by D. Archer [et al.]. UCREL Technical Papers Series, Volume 16. Lancaster, U.K.: UCREL, Computing Department, University of Lancaster.
- Harris, Randy A. 1993. *The Linguistics Wars*. New York: Oxford University Press.
- Hasnain, S. Imtiaz. 1985. Standardization and Modernization of Languages: A Case Study of Urdu. Ph.D. diss., Jawaharlal Nehru University.
- Humayoun, Muhammad, Harald Hammarström, and Arne Ranta. 2007. Implementing Urdu Grammar as Open Source Software. Paper presented at the Conference on Language and Technology, University of Peshawar 7–11 August. [<http://www.lama.univ-savoie.fr/~humayoun/UrduMorph/downloads/UrduExtendedAbstract.pdf>].
- Husain, Intizār. 2000. “Urdū kā Tahzībī Mizāj” [Urdu’s Cultural Temperament]. *The Annual of Urdu Studies* 15(2):372–76. [Originally published in *Akbbār-e Urdū* (Islamabad) May 1998.]
- Husain, Syed Afaq and Syed Hassan Amin. 2002. “A Multi-Tier Holistic Approach for Urdu Nastaliq Recognition.” In *International Multi Topic Conference 2002: Abstracts (December 27–28, 2002 at Hotel Marriott and Karachi Institute of Information Technology)*. Piscataway, N.J. Institute of Electrical and Electronics Engineers. [Full-text of papers available on CD-ROM version].

- Hussain, Sarmad. 1997. *Phonetic Correlates of Lexical Stress in Urdu*. Ph.D. diss., Northwestern University.
- Ijaz, Madina and Sarmad Hussain. 2007. "Corpus Based Urdu Lexicon Development. In *The Proceedings of Conference on Language Technology (CLT07)*. Peshawar: University of Peshawar, Department of Computer Science.
- Iqbal, Jonaid. 2010. "Urdu as Official Language Sought." *Dawn* (Karachi), 26 May.
- Jakobson, Roman. 1960. "Closing Statement: Linguistics and Poetics." In *Style in Language*. Ed. by Thomas A. Sebeok. Cambridge: M.I.T Press.
- Jinnah, Mohammad Ali. 1989. *Quaid-i-Azam Mohammad Ali Jinnah: Speeches and Statements as Governor General of Pakistan 1947-48*. Islamabad: Ministry of Information & Broadcasting, Government of Pakistan.
- Kachru, Yamuna. 1988. "Writing in Hindi and English." In *Writing Across Languages and Cultures: Issues in Contrastive Rhetoric*. Ed. by A.C. Purves. Newbury Park (CA): Sage Publications.
- Khan, Iqtidar H. 1999. *A Contrastive and Comparative Study of Standard Urdu and Standard Hindi*. Aligarh: Aligarh Muslim University Press.
- \_\_\_\_\_. 1989. *Studies in Contrastive Analysis*. Aligarh: Aligarh Muslim University, Department of Linguistics.
- and K.S. Mustafa. 1984. "Simple Finite Verbs in Dakhini." *Indian Linguistics* 45:1-4.
- Khan, Masud Husain. 1969. "Urdu." In *Current Trends in Linguistics, Vol. 5, South Asia*. Ed. by Thomas A. Sebeok. Mouton: The Hague.
- Khan, M.A. [et al.]. 2007. "Treatment of Pronominal Anaphoric Devices in Urdu Discourse." In *Proceedings of the International Conference on Emerging Technologies, 2006*. Peshawar: Department of Computer Science, Peshawar University.
- Khan, S.A. and M.A. Khan. 2009. "Computation of Gender of Urdu Nouns." In *Proceedings of the Conference on Language and Technology 2009*. [http://www.crupl.org/clt09/download/Papers/Paper10.pdf].
- Khaṭṭak, Parēshān. 1987. *Lisānī Rābiṭa: Urdū, Sindhī, Pashtō, Panjābī aur Baločī kē Mushtarak Alfāz* [Linguistic Relationship: Words Common to Urdu, Sindhi, Pashto, Panjabi and Balochi]. Islamabad: National Language Authority.
- Khurshid, Salman. 2006. Preface to *Redefining Urdu Politics in India*, by Ather Farouqui. New Delhi: Oxford University Press.
- Kiparsky, Paul. 1973. "The Role of Linguistics in a Theory of Poetry." *Daedalus* 102: 231-44.
- LaDousa, Chaise. 2010. "On Mother and Other Tongues: Sociolinguistics, Schools, and Language Ideology in Northern India." *Language Sciences* 32(6):602-14.
- Lelyveld, David. 1993. "Colonial Knowledge and the Fate of Hindustani." *Comparative Studies in Society and History* 35(4):665-82.
- Linguistic Society of India. 2008. [Report of the 30th All India Conference and the Golden Jubilee Celebrations of Post-Graduate Teaching of Linguistics at the Deccan College, held at the Department of Linguistics, Deccan College Post-Graduate & Research Institute from 26-28 November.] [http://www.lsi.org.in/AICL\_Report.aspx].

- Lodhi, Saeed. 2004. Robust Urdu Character Recognition Using Fourier Descriptors. Ph.D. diss., University of Denver.
- McEnery, Anthony [et al.]. 2000. "EMILLE: Towards a Corpus of South Asian Languages." *British Computing Society Machine Translation Specialist Group* 11:1–9.
- Metcalfe, Barbara D. 2006. "Urdu in India in the Twenty-First Century." In *Redefining Urdu Politics in India*. Ed. by Ather Farouqui. New Delhi: Oxford University Press.
- Mufwene, Salikoko S. 2008. *Language Evolution: Contact, Competition and Change*. London; New York: Continuum.
- \_\_\_\_\_. 1998. "Native Speaker, Proficient Speaker and Norms." In *The Native Speaker: Multilingual Perspectives*. Ed. by Rajendra Singh. Language and Development Series, Vol. 4. New Delhi; London: Sage Publications.
- Mukhtar, Omar, Srirangaraj Setlur, and Venu Govindaraju. 2009. "Experiments on Urdu Text Recognition." In *Guide to OCR for Indic Scripts: Document Recognition and Retrieval*. Ed. by V. Govindaraju and S. Setlur. Advances in Pattern Recognition Series. London: Springer-Verlag.
- Mustafa, Khateeb S. 2008. Urdu Broken Plurals: A Morpho-Semantic Study. Paper presented at the 30th All-India Conference of Linguists, Pune. 20–28 November.
- \_\_\_\_\_. 2000. *A Descriptive Grammar of Dakkhini*. New Delhi: Munshiram Manoharlal.
- Nāhīd, Kishvar. 2010. "Pākistān Mēn Shu'ba-e Lisāniyāt Kahīn Nahīn Hai" [Nowhere in Pakistan is There a Department of Linguistics]. *Jañg* (Karachi) 12 July.
- Naim, C.M. 1957. "Urdū kē Kūzī Tajšautiyē" [The Retroflex Phonemes of Urdu]. *Urdū Adab* (June):31–38.
- \_\_\_\_\_, trans. 1956. "Urdū aur Hindī kā Talaffuz." *Urdū Adab* (December):90–102. [Translation of Bailey, T. Grahame. 1922. "Pronunciation of Urdu and Hindi." *Bulletin of the School of Oriental Studies* 2(3):539–47.]
- Narang, Gopi Chand. 1961. *Karkhandāri Dialect of Delhi Urdu*. New Delhi: Caxton Press.
- Nardella, Umberto. 2008. *Glossary of Hindi/Urdu and English Linguistic Terminology*. New Delhi: Star Publications.
- Onions, C.T. 1955. *The Oxford Universal Dictionary on Historical Principles*. Prepared by William Little, H.W. Fowler, J. Coulson. 3rd edition, corrected and with revised addenda. Oxford: The Clarendon Press.
- Palmer, F.R. ed. 1968. *Selected Papers of J.R. Firth 1952–59*. Bloomington; London: Indiana University Press.
- Parekh, Rauf. 2009. "Urdu Linguistics and Abul-Lais Siddiqi." *Dawn* (Karachi) 7 September.
- Platts, John T. 1884. *A Dictionary of Urdu, Classical Hindi, and English*. 4th impression. London: Crosby Lockwood and Son, 1911.
- Prew, Martin. 2011. "Matric Results Do Make Sense." *Mail & Guardian*. 14 Jan. [Online ed. <http://www.mg.co.za/article/2011-01-14-matric-results-do-make-sense>].
- Qadri, S.G. Mohiuddin ("Zor"). 1930. *Hindustani Phonetics: A Phonetic Study of*

- Hindustani Language as Spoken by an Educated Person of Hyderabad, Dn.* Hyderabad Deccan, India: Maktaba-e Ibrāhīmiya.
- Qureshi, Bashir Ahmad. 1971. *Kitabistan's 20th-Century Standard Dictionary, Urdu into English*. Lahore: Kitabistan Publishing Co.
- Rahman, Tariq. 2010. *Language Policy, Identity, and Religion: Aspects of the Civilization of the Muslims of Pakistan and North India*. Islamabad: National Institute of Pakistan Studies, Quaid-i-Azam University.
- . 2002. *Language, Ideology and Power: Language Learning Among the Muslims of Pakistan and North India*. Karachi: Oxford University Press.
- . 1998. "Regional Report on Pakistan: Indo-European." In *The Yearbook of South Asian Languages and Linguistics 1998*. Ed. by Rajendra Singh. New Delhi; London: Sage Publications.
- Rampton, M.B.H. 1990. "Displacing the 'Native Speaker': Expertise, Affiliation, and Inheritance." *English Language Teaching Journal* 44(2):97–101.
- Rauf, Zia. 1997. *Phonology of Delhi Urdu*. Delhi: Shipra Publications.
- Rehman, M.A.U. 2010. "A New Scale Invariant Optimized Chain Code for Nastaliq Character Representation." In *Computer Modeling and Simulation, 2010, Vol. 4*. Proceedings of the Second International Conference on Computer Modeling and Simulation (ICCMS) Sanya, China, 22–24 January. New York: Institute of Electrical and Electronic Engineers (IEEE), Catalog Number CFP1097F-PRT.
- Rizvi, Syed Muhammad Jafar. 2007. *Development of Algorithms and Computational Grammar for Urdu*. Nilore, Pakistan: Pakistan Institute of Engineering and Applied Sciences. [[http://www.lingfest.arts.usyd.edu.au/LFG/rizvi-j\\_10.pdf](http://www.lingfest.arts.usyd.edu.au/LFG/rizvi-j_10.pdf)].
- Sabzvārī, M.A. 1995. *Iṣṭilāḥāt-e 'Imrānīyāt* [Sociological Terminology]. Islamabad: Muqtadira Qaumī Zabān [National Language Authority].
- Saleemi, Anjum P. 2004. "Case and Agreement in Hindi-Urdu." In *The Yearbook of South Asian Languages and Linguistics 2003*. Ed. by Rajendra Singh. Berlin; New York: Mouton de Gruyter.
- . 1994a. "On the Acquisition of Split-Ergativity: Some Evidence from Urdu." In *The Proceedings of the Twenty-Sixth Child Language Research Forum*. Ed. by Eve V. Clark. Stanford: Center for the Study of Language and Information.
- . 1994b. "If That's the Case, I Don't Agree: Case, Agreement and Phrase Structure in Hindi-Urdu." In *Papers from the Fifteenth South Asian Language Analysis Roundtable Conference 1993 (SALA 15)*. Ed. by Alice Davison and Frederick M. Smith. Iowa City, IA: South Asian Studies Program, University of Iowa.
- . 1992. *Universal Grammar and Language Learnability*. Cambridge: Cambridge University Press.
- Sattar, Sohail A. [et al.]. 2008. "Implementation Challenges for Nastaliq Character Recognition." In *Wireless Networks, Information Processing and Systems*. Communications in Computer and Information Science, 20. Ed. by D.M.A. Hussain [et al.]. Berlin; Heidelberg: Springer-Verlag.

- Schmidt, Ruth Laila. 1981. *Dakhini Urdu: History and Structure*. New Delhi: Bahri Publications.
- Shukla, Shailendra. 2008. Phonological Error by Hindi Speakers While Learning Urdu. Paper presented at the 30th All-India Conference of Linguists, Pune. 20–28 November.
- Tuite, Keven, Asif Agha, and R. Graczyk. 1985. "Agentivity, Transitivity and the Question of Active Typology." In *Papers from the Parasession on Causatives and Agentivity at the Twenty-First Regional Meeting*. Ed. by William H. Eilfort, Paul D. Kroeber and Karen L. Peterson. Chicago: Chicago Linguistic Society.
- Wagner, Michael and Katherine McCurdy. 2010. "Poetic Rhyme Reflects Cross-Linguistic Differences in Information Structure." *Cognition* 117(2):166–75.
- Watkins, Calvert. 1995. *How to Kill a Dragon: Aspects of Indo-European Poetics*. New York; Oxford: Oxford University Press.
- Yaqin, Amina. 2006. "The Communalization and Disintegration of Urdu." In *Redefining Urdu Politics in India*. Ed. by Ather Farouqui. New Delhi: Oxford University Press.