

CURRENT PERSPECTIVES IN URDU LINGUISTICS

Khateeb S. Mustafa

If one is to understand the term 'linguistics' in its non-technical sense one may argue that linguistic activity in Urdu quite old. All those attempts of Urdu scholars on the one hand to purge and veer away the language from Persian influence and on the other to codify and fix its idiom may be labelled as some sort of linguistics. In this regard notable are the efforts by Insha Allah Khan Insha and Shah Hatim. Insha's work 'Darya-e-latafat' wherein he discusses and sets norms about the use of language may be dubbed as an important linguistic activity.

Today however we understand linguistics in its technical sense and label the discipline as 'modern linguistics'. Work from this view-point in Urdu was commenced as early as 1930 with the publication of Hindustani Phonetics by S. M. Qadiri Zor. After Zor however there was a brief lull. The thread left by Dr. Zor was picked up by many scholars in and after fifties and a good volume of literature on Urdu linguistics came forth. However much of the work relates to Dakkhini Urdu, a variety of Urdu cultivated and spoken in southern part of India. That way, today what we call as 'standard Urdu' is only second to Dakkhini Urdu in this respect. It may also be argued that compared to other Indian languages, say like Tamil, Telugu, Hindi etc work on standard Urdu from the perspective of modern linguistic theory is very meagre and insignificant. The area is almost barren. The present paper takes a look at the situation and reflects on a few areas that may be relevant for the current and future research on Urdu from the view-point of modern linguistic theory.

1. DESCRIPTIVE ANALYSES ON STANDARD URDU

Urdu has a rich past. In pre-independent India it held a pre-eminent position in many spheres of social life viz., literature, mass-media, a medium of instruction in education etc. However after independence with the re-organization of Indian states it became a stateless language. Further in states it was not a major language. Therefore it came to be a 'nobody's baby'. However despite many odds the language proved to be a living language by way of the demography of its speakers and its use in cultural and social life of India. Therefore the governments at the central and

IJL (Interdisciplinary Journal of Linguistics) Vol(2), University of Kashmir.

provincial levels realised its strength and significance. As a result many new Urdu departments were established in both central and provincial level universities. Similarly many Urdu academies at the provincial level, an Urdu council and a Central Urdu University at the national level were also established. Here the aim was to sustain and promote the growth of the language. No doubt these depts., academies and other bodies are engaged in maintaining and promoting the language in their own way. But almost all of the work they are carrying out is on literary lines. The present author has observed that they do not have any idea about the 'linguistics' of the language. Therefore one cannot expect of them to think on any work on the basis of theory of modern linguistics. Be that as it may, a few departments of linguistics that are established in the state and central universities focus mainly on the language of the state or on some major language or on some theoretical issue in linguistics. As a consequence of all this Urdu remained and continues to be 'a neglected child' as far as the work from the view-point of modern linguistic theory is concerned.

As stated above a good and fair amount of work has been done on languages like Tamil, Telugu, Hindi, Kashmiri from the view-point of descriptive linguistic theory. Barring a few stray works no serious analysis of standard Urdu on these lines is available. Therefore the author feels that this kind of work with respect to standard Urdu has a great relevance even today.

One may argue that linguistics is much advanced today and that it has come a long way since the days of Bloomfieldian linguistics and that currently we are thinking in terms of 'computer and language', 'brain and language' etc. Nevertheless the basic linguistics has its own relevance and more so in relation to standard Urdu as serious analyses have not emerged. Therefore analysis of standard Urdu on descriptive lines is a strongly recommended area in the current scenario. But how to go about it? As is generally followed in linguistics we may locate the area of spoken standard Urdu. Theoretically considering the Urdu spoken in the area as a homogenous body we may go about collecting samples of spoken language and subject the data for analysis. We may work out the phonetics and phonology of the language and produce descriptions on these areas. While working here one may not restrict oneself to only Bloomfieldian model but consider other approaches and schools of linguistic thought. The emerging phonetic and phonological analyses have wide-ranging applications in the current expanding scenario of the language. The findings may serve as reference points not only in language pedagogy but also for the language used in mass and entertainment media.

Similarly studies on morphology, syntax and semantics are equally important. Due to many cross linguistic trends, the present Urdu is undergoing many changes in its morphology, syntax and semantics. Thus for example with the pronoun *ap*, –o ending imperative form is slowly gaining currency. We still are not sure enough

whether such standards of usage are acceptable in current standard Urdu. Likewise there are many nouns whose gender has been an issue of debate. The occurrence of genitive phrases and noun phrases modelled on Persian pattern is a usual phenomenon in standard Urdu. These phrases seem to have been fully absorbed and assimilated. The morphological and syntactical rules of these are yet to be fully worked out and formulated. At the semantic level too there are many issues which are needed to be studied. Thus for example many words borrowed into Urdu from Persian and Arabic have undergone changes not only in their grammatical status but in their meaning too—that is, the latter is either narrowed down, expanded or altogether changed. For instance /əoqat/ ‘timings’, broken plural of /vaqt/ ‘time’ (Arabic) can also be used in Urdu as a singular noun with a meaning of ‘status, position, capacity’. Thus /meri Itni əoqat kəhā ke məĩ do do nəukər rəkh səkū/ ‘I am incapable of having two servants’. All this is to show that work on morphology, syntax and semantics on standard Urdu is important and useful. While working on such studies we should not lose sight of various theories and approaches of morphological and syntactic processing. Thus it would be worthwhile to study and analyse standard Urdu from the standpoint of theory of transformational grammar or GB theory of syntax as except a few papers not much is available in the area. To sum up it may be said that work from these perspectives is of vital importance as far as standard Urdu is concerned.

2. SOCIO-LINGUISTIC STUDIES ON URDU

As stated above today Urdu is a stateless language. But it exists almost in whole of India in the form of its islands. This peculiar linguistic situation has given rise to different varieties of Urdu. Thus one may recognize varieties such as Mumbaiyah Urdu, Kolkatta Urdu, Bhopali Urdu, Hyderabad Urdu etc. In each of these areas the language is converging towards the local language and in some instances it is also making an impact on the respective local language. What is meant here is that the situation has given rise to good scope for convergence studies. Along with the studies on convergence we may also undertake studies on socio-linguistic phenomena such as bilingualism, code-mixing and code-switching etc with respect to the contemporary Urdu. The findings of such studies may provide good insights into language use and language maintenance in different spheres of social life at the individual and societal level with respect to Urdu.

Along with the above, attitudinal studies with respect to Urdu speakers would be a useful area of research. This is because there is a general feeling that the base of Urdu medium education is gradually shrinking. More and more Urdu children are looking towards English or the regional language for their school study. On the other side the language is surging ahead with its more and more use in mass and entertainment media. May be, an instance is the popularity of Urdu ghazal; one may

see it is liked by all and sundry. Studies focusing on attitudes of the Urdu speakers towards their own language and attitudes of others towards Urdu probably will provide answers to such contradictory phenomenon. The findings of such studies will have important implications in the area of education too.

Generation gap with respect to the reading and writing skills of Urdu would form one more fruitful area of research in Urdu socio-linguistics. To illustrate the point one may give the example of linguistic scenario vis-à-vis Urdu in states that are generally known as 'Hindi states' in media. It has been observed that in these states there is a big gap between older and younger generation with respect to the knowledge of Urdu, especially in reading and writing skills. As a contrast down south in states viz., Andhra Pradesh, Maharashtra, Karnataka the gap seems to be insignificant due to a fairly good state of Urdu education at school level. In many of these 'Hindi states' there is a sizeable strength of Urdu speakers. But one may not find Urdu medium govt schools in proportion to the number of speakers of the language. The result is the tradition of Urdu learning through govt. schools in the early age is almost lost. It is also a fact that in many of these states the instrumental value of Hindi education is being perceived as higher (in comparison to Urdu education) by the Urdu speakers. Further the structural similarity between Hindi and Urdu and the ease with which one can shift from Urdu to Hindi at basic level seems to be one of the reasons why Urdu speakers in these states adopt to Hindi education. The net result of all these factors is that almost the entire younger generation of Urdu speakers in states like U. P., Delhi, Rajasthan, Madhya Pradesh etc is ignorant of reading and writing Urdu. Therefore a study in this area may have implications not only for the improvement of status of Urdu education but in the maintenance of language too.

The use of Urdu in advertisements, Hindi T. V. serials, films and newspapers is another potential area of research in Urdu socio-linguistics. Looking at the language that is being used on the above media one feels that here, a lot of code-mixing is taking place from Urdu to Hindi. Mustafa (2006) in one of his studies on Urdu elements in the current Hindi of Hindi newspapers has pointed out that in recent times Hindi newspapers are adopting more and more Urdu words and phrases. He further points out that the earlier trend of highly sanskritized Hindi which was the norm in mass media has almost waned. Many such studies on language use in different spheres with respect to Urdu may be taken up as the issues have implication for the identity of the language.

One more potential area of research on Urdu may be that of the diglossia like situation that has come to exist in the language. To elaborate the point Urdu, as mentioned above, exists in almost all regions of India in the form of islands amidst the major languages. As stated above one may find many varieties of spoken Urdu but when it comes to written Urdu, especially the literary and journalistic variety,

irrespective of the region similar grammatical norms are noticed. The author had occasion to listen to poets from various regions of the country and also read newspapers published from different cities of India in east, west, north and south. The language used in these poetic compositions and newspapers has almost similar structural characteristics. If there was any divergence it pertained to the vocabulary used and this was not very significant. In our view the situation here is almost akin to that of Arabic where the norms for written language are well laid out. A speaker of Arabic from any region is familiar with and follows these norms. More investigation into the characteristics of written Urdu in different regions of the country will unravel the kind of diglossia we have here. The study can be useful to compare and contrast the norms of written Urdu in Indian situation with those of written Urdu in other regions of the world like Pakistan, Britain, Dubai, Kuwait, Mauritius etc.

3. COMPUTATIONAL LINGUISTICS AND URDU

Computational linguistics is a recent discipline and many novel studies are being pursued in this area in relation to different Indian languages. Though Urdu has retained its character as one of the leading literary languages it has very little to be proud of as far as research into computational linguistics (vis-à-vis Urdu) is concerned. Barring a couple of word processors literally nothing is available. Therefore the task here is quite stupendous. Keeping the current needs of the language and current trends in computational linguistics the foremost need seems to be that of good word processors. Word processors are useful programmes to carry out day to day activities in the original script of the concerned language and also for desktop publishing. Presently, in Urdu, we have only two wordprocessors viz., InPage Urdu and Nafees Nastaliq that are seen being used widely. InPage was developed in India at Delhi where as Nafees Nastaliq is a product of Centre for Research in Urdu Language Processing (CRULP), Lahore, Pakistan respectively. Both are window-based programmes and have a variety of fonts. Further they have facility to prepare a document in nastaliq and naskh both. In Indian situation all the publishing work is being carried out in InPage Urdu and that way it is a popular programme in the arena of Urdu desktop publishing. However it has certain drawbacks. Thus for example being a ligature-based programme one cannot type the individual allographic shapes of alphabetical letters which are sometimes needed in language pedagogy. Further it is not possible to have clearly demarcated word boundaries with equal space between different words. There is no facility of transliteration into roman or any other Indian language. Further the word-processed data cannot be transported from InPage Urdu to other applications easily. No electronic communication such as sending and receiving emails, attachments is possible. No ASCII or UNICODE code is assigned to the alphabets. Above all one needs to purchase InPage Urdu at exorbitant price.

Nafees Nastaliq as said above is also a window-based programme. Though Unicode is assigned to the alphabets the word-processed text seems to lack aesthetic beauty and visual appeal that one finds in Urdu calligraphy.

Software tools for word count, letter count, frequency count and the like are lacking in both the softwares. In view of all this there is an urgent need for the research into and development of good word processors in Urdu.

Electronic dictionaries: A good, fairly voluminous monolingual Urdu dictionary is need of the hour. This dictionary will be of multiple use and can function as a foundation for other dictionaries, viz., bilingual and trilingual dictionaries; dictionary of synonyms, antonyms, idioms, proverbs etc., standard spelling and standard pronunciation dictionaries.

Other areas: In addition to the above research into other NLP areas like development of OCR systems, morphological analysers, syntactic parsers, machine aided/machine translation system, text to speech systems in relation to Urdu need to be planned and undertaken so that the language may grow in this direction too and may catch up with other Indian languages.

REFERENCES

Fatihi, A.R. 1991. *The Language of Advertising and TV Commercials*. New Delhi: Bahri Publications.

Mirza, Shaukat, A. and Fatima, Sameen, S. 2003. *Urdu Informatics: Challenges and Opportunities*. Keynote Address presented at National Meet on Urdu Language Technology : Prospects and Problems, University of Hyderabad.

Mustafa, Khateeb S. 2000. *A Descriptive Grammar of Dakkhini*. New Delhi: Munshiram Manoharlal Publishers Pvt. Ltd.

Mustafa, Khateeb S. 2000. *Dynamics of Code-mixing: The Case of Urdu in TV Commercials*. An unpublished paper presented in the Seminar on 'Innovations in the language of News Media', AMU, Aligarh.

Mustafa, Khateeb S. 2003. *Computerization of Urdu: The Task Ahead*. An unpublished paper presented at National Meet on Urdu Language Technology : Prospects and Problems, University of Hyderabad.