This chapter gives a general introduction about Kashmiri, its language groups, dialects and varieties, nature of sounds and Kashmiri phonetics and phonology.

Kashmiri, also called ‘kaːʃur’ or ‘kaːʃir zabːɔːn’ is primarily spoken in the Kashmir province of Jammu and Kashmir state in India, and in other parts of world where Kashmiris are settled. The state of Jammu and Kashmir is the northernmost state of India and shares its boundary with Pakistan, China and Afghanistan. This puts Kashmiri in a position of being in the context of several major language groups, some of them being: Tibeto-Burman languages such as Ladakhi, Balti and Gilgiti; Indo-Aryan languages such as Pahari languages, Dogri; and the language ‘isolate’, Burushaski.

Linguistically, Jammu and Kashmir state and its surrounding areas constitute a complex region in the sub-continent and its linguistic complexity is evident by the segmentation of population into the following language inventories:

1. Dard group which includes within it the languages/dialects such as Kashmiri, Shina, Kashtiwari, Poguli, Siraji, Rambani and Kohistani (Grierson, 1919; Kachru, 1969).
2. Indo–Aryan language family which includes the languages like Punjabi, Urdu, Hindi, Dogri, Lahanda, Pahari varieties and Gojri.
3. Tibeto–Burman family which includes languages such as Ladakhi and its dialects.
4. Pashtu (Eastern Iranian language) speakers, immigrants from Afghanistan and the North-West Frontier, and some part of the Central Asia region. The major
concentration of these speakers in Kashmir is at Gotlibagh in Ganderbal (District Ganderbal) and Wantrag in Anantnag (District Anantnag).

5. Hitherto linguistically unexplored or underexplored speeches like that of Sheikhs/Watals.

1.1 Linguistic Affinity

The position of Kashmiri vis-à-vis the Indo-Aryan language family is controversial. Some linguistic scholars consider Kashmiri as an Indo-Aryan language while another group places it under the Dardic branch of the Aryan group. Wali and Koul (1997) name Prof. Kuhn as the first to suggest (*The Indian Antiquary*, May 1887) that the Hindukush dialects together with Kashmiri formed a separate group within the body of the Indo-Aryan languages and suggested distraction viz-a-viz linguistic features especially based on phonetic peculiarities.

It is not easy to assert the linguistic affinity of Kashmiri with the Indo-Aryan family because of its several unique characteristics, which differentiates it from other Indo-Aryan languages like Punjabi and Hindi. Grierson (1906) lists Kashmiri as a Pisacha language. He says that Nagas originally inhabited Kashmir and because of Kashypa, Pisachas could settle there. These Pisachas used to speak a language, which was closer to Shina (a Dard language) before they came under the Aryan influence. Grierson (1915) concludes that Kashmiri is a mixed language having at its base a language of Dard group of Pisacha family allied to the language, Shina. According to him, Kashmiri is highly influenced by the Indian culture and literature and a large part of its vocabulary has an Indian origin. He however claims that its etymology, phonology, syntax and prosody is Pisacha and states that it must be classed as such, and not as a Sanskrit form of speech. In his later writings, Grierson preferred the term 'Dardic'.

Grierson assumed that in the pre-historic times there was a language known as Aryan, spoken by the common ancestors of the Iranians and the Indo-Aryans in the Oasis of Khiva ‘the original home whence the Aryans separated’. In his linguistic survey of India, Grierson (1919) shows bifurcation of the Aryan stage into three distinct sister groups.
Morgenstierne (1947, 1961) rejected Grierson’s view and claimed Dardic language to be of pure Indo-Aryan origin though Dardic had preserved many archaisms which were lost in later Indo Aryan languages. He concluded that Aryan is subdivided into Indo Aryan and Eranian only, and Khowar and Dard are included in Indo Aryan whereas Kafir is placed under Eranian.

He further called Kishtwari, a true dialect of Kashmiri.

Chatterji (1963) states that Kashmiri, at least in terms of its base, belongs to the Dardic group of Aryan or Indo-Iranian. He believes that a group of the Aryan who came before 1000 BC and spoke dialects akin to Rig Veda though with some special characteristics settled in Kashmir. He then goes on to state that

Another wave of Aryans, with their Indo-Aryan "spoken" Sanskrit (and subsequently Prakrits), came and settled down in Kashmir and other Himalayan areas ... in this way, Kashmir, in spite of a Dardic
substratum in its people and its speech became part of a Sanskrit culture world of India. The Indo-Aryan Prakrits and Apbbramsha from the Midland and from Northern Punjab profoundly modified the Dardic bases of Kashmiri, so that one might say that Kashmiri language is a result of a very large overlaying of a Dardic base with Indo-Aryan elements. (Chatterji 1963:256)

Wali and Koul (1997) mention that these features include phonetics, phonology, morphology, pronominal system, syntax and prosody. Many other features like insertion of epenthetic vowels, aspiration of final unvoiced stops, absence of voiced aspirates, three way distinction of pronominal system are different from those in Indo-Aryan languages. Zakharyin (1984) holds the view that Kashmiri is a member of Indo-Aryan family. According to him, The more we learn about the Dardic languages.... the more evident it becomes that G.A.Grierson might have been wrong to separate Kashmiri from the Indo-Aryan language stock, and that perhaps J.Bloch (1934) was right in stating that Kashmiri only primordially has been Dardic and later underwent a heavy ‘Indo-Aryanization’. (1984:29)

Koul and Schmidt (1983) show the characteristics of the Dard group in terms of phonological and morphological features in Kashmiri and Shina. In contrast to Shina, Kashmiri has developed a system of palatalized consonants and more importantly, a verb-second order, which is typical of Kashmiri and Kishtwari, which is a dialect of Kashmiri.

Handoo (1994) argues that linguistic affinities indicate that Kashmiris formed a branch of the race, which brought the language of the Indo-Aryans to India. The dialects spoken in Baltistan, Dardistan and Kashmir valley are closely related, belonging to the Dardic group of the family. He further quotes Grierson and indicates that Kashmiri has a Dard base and is one of the independent (the others being Indo-Aryan and Iranian) branches of Indo-European family.

Wali and Koul (1997) also suggest that Kashmiri belongs to the Indo-Aryan hill language family. According to them Kashmiri and related hill languages are called Dardic giving an impression that they form some sort of a separate branch from Indo-Aryan. They also take into consideration the individuality of Kashmiri as the
only language of Dardic group that has a rich literary tradition dating from the thirteenth century onwards.

It appears that Dardic group, though small as compared to Iranian and Indo-Aryan groups has a separate identity and possess its own peculiarities. Of these Dardic languages Kashmiri is the only language reaching literary civilization. Even today some of the basic terms like earlier numerals and kinship terms are close to Shina words and thereby have a Pisacha origin. Pisacha, according to Grierson (1906), was the cover term used to refer to Dardic languages. Furthermore, Grierson (1906) also points out that its consonants, vowels and pronunciation appear to have a Dard origin. Also, he claims that on phonological and morphological considerations, its base, phonetic system, syntax and prosody is associated with Pisacha. Local tradition as pointed out by historians like Sufi (1996), Rabbani (1981), and others also hold that the first permanent settlers in the valley, Nagas were outnumbered by Dards (Pisachas) from the north long before the Mauryan age. The Pisachas were followed by Aryans from Central Asia and across time Kashmir became a celebrated centre of Sanskrit, which had its influence on Kashmiri. This could have resulted in Aryanising Kashmiri. It can again be stated here that the position of Kashmiri is controversial and it is beyond the scope of the topic to dwell more on this issue.

1.2 Dialects and Varieties

Grierson (1919) claims that Kashmiri has only one true dialect i.e. Kishtwari and a number of mixed dialects like Poguli, Siraji and Rambani and in the farther east of Riasi, there are some mixed dialects but not much is known about them except that they are mixed varieties of Kashmiri and Chibhali form of Lahanda. Kishtwari is spoken in the Kishtwar region, Siraji-Kashmiri is spoken in some rural areas of Doda and Ramban tehsil which is under deep influence of other Pahari speeches. Poguli is a dialect spoken generally by the inhabitants of Pogul region of the Ramban tehsil. In the valley, the Kashmiri speaking area is ethno-semantically divided into three groups viz: maraːz- spoken in southern and south-eastern region. kamraːz- spoken in northern and north western region. yamraːz- spoken in Srinagar and its neighbouring areas.
There are some minor linguistic variations between these varieties. The main variations are phonological and at times semantic (Koul 2001). Some of the main characteristics of these speech variations are as follows.

1. The Marazi Kashmiri has the flap /ṛ/ which is represented by /r/ in Kashmiri of Kamraz area and Srinagar.

2. The Marazi Kashmiri has the progressive or indefinite aspect suffix -‘an’ with the verb roots, while –‘a:n’ is used instead in other two varieties.

3. The Kamrazi Kashmiri distinguishes itself from the variety spoken in the Maraz as well as in Srinagar mainly by the use of ‘characteristic’ intonation and stress.

4. A number of vocabulary items are different in Kashmiri spoken in the above three regions.

1.3 Nature of Kashmiri sounds

Kashmiri shares several phonetic features with other Indo-European languages of the area. It also shares some of the specific characteristics of a group containing languages like Shina, Poguli.

Some of these interesting features are presence of contrastive palatalization, co-occurring with the large set of consonants, presence of full fledged central vowel system, absence of contrastive voiced aspirates and insertion of epenthetic vowels. /i/, /i:/, /a/, /a:/, presence of an indefinite article, and a three way distinction in the pronominal and past tense system. Also, in Kashmiri the verb always occurs in the second position in a finite clause excepting the relative constructions. The word order in Kashmiri, thus resembles the one in German, Dutch, Icelandic, Yiddish and a few other languages. These languages form a distinct set and are currently known as Verb Second (V-2) languages. From a morphological point of view, Kashmiri is a mix of agglutinating and inflectional type (Wali and Koul 1997). It shows both types of morphemic processes across most lexical categories such as nouns, verbs, and adjectives. In other words, one can say Kashmiri is morphologically very rich.

Several eminent linguists, both native speakers of Kashmiri as well as others described the sound system of the language either in research papers or in book-size
publications (such as Kelkar and Trisal 1964; Bhaskararao et al., 2009). They have included a good amount of discussion based upon perceptual and/or instrumental phonetic findings.

1.4 Word level Phonetics and Phonology

The present dissertation begins with the review of the phonetics of segmental phonemes that occur at word level. This review is based on the findings detailed in some of the important publications on the topic, coupled with the author’s fresh and critical approach. The author being a native speaker of standard Kashmiri shares her speech patterns with the majority of Kashmiri speakers generally from Srinagar. Such a re-look at Kashmiri phonology has resulted in some interesting phenomena such as the discovery of a contrast between a voiced labiodental fricative /v/ and the conventional bilabial approximant /w/, [a] and [a•] as two freely varying allophones of /a/, [ua] as a conditioned variant of /o:/, [ia] as a free variant of /a:/, [e] an allophone of /e/, etc.

It is pertinent to point here that language in actual use is very different from the simplistic sound-after-sound and word-after-word perfect-looking purely segmented entity. In continuously spoken utterances, use of language involves overlappings of sounds and words, and these have all been neglected in linguistic studies especially when it came to Kashmiri. The work done on Kashmiri so far is based only on the competence aspect (in its perfect non-overlapping mode) and practically no work has been done taking the performance side into consideration. This work is the first attempt to look at Kashmiri phonetics and phonology from the ‘performance-generated’ spoken perspective rather than the assumed 'perfectly' segmented, non-combinatorial and non-inter-influencing perspective.

1.5 Phrase level Phonetics and Phonology

Although, phoneticians have always recognized that sounds of a language set up individual word-level changes, within the last decade there has been a renewed and concentrated effort to study the sound system of languages based on units of utterance that are larger than lexical words, such as phrases and sentences. This interest has several underlying reasons such as applications to language technology, applications to speech pathology, etc.
The International Phonetic Association sponsored a special international symposium on ‘Patterns of Production-Perception-Phonology’ in Germany in October 2000. Selected papers from this symposium were published as a part of the 31st volume of the JIPA (Journal of International Phonetic Association) as Kohler (2001) in his presidential foreword pointed out one of the ‘hot’ issues in Phonetics is “Phrase-level phonetics”.

Kohler (2001) hoped that the special publication “will trigger on increased interest in Phrase level Phonetics and Phonology in a great number and variety of languages and thus complement the WORD-phonemic/allophonic language illustrations, which have become a standard feature of the Journal, by more comprehensive phonetic accounts of communication in individual languages, on the one hand, and in human language in general”. Kohler (2001) concludes his foreword saying that the special papers in the volume “converge in a realization of the need for more intensive research into patterns of speech sounds in unscripted communication”.

The need for study of non-canonical pronunciation of forms that occur at the phrase-level as well as higher levels has been constantly felt by phoneticians and speech scientists as recently as July 2011. When the Journal of Phonetics released a special issue in its volume 39 on the “Phenomenon of Reduced Pronunciation Variants”. This issue incorporates eight research papers that were presented in a workshop earlier. The common thread of these papers was the broad conclusion “about what kind of acoustic information humans produce in reduced speech, and how listeners use that information to retrieve the speaker’s message if not all of the speaker’s intended segments or words” (Ernestus and Warner 2011:259).

Ernestus and Warner (2011:256) clarify that since ‘speakers/listeners comprehend reduced pronunciation variants, the processing of these variants must be accounted for by psycholinguistic models of speech comprehension and production. The existing models were developed mostly from evidence about the processing of words produced carefully in isolation or in simple sentences in the laboratory.

The pronunciation of a word ‘produced carefully in isolation’ can be considered as its ‘canonical pronunciation’. Such canonical pronunciation is ‘stored in
the mental lexicon in the form of abstract representations’ (Ernest and Warner 2011:256).

Following Goldinger (1998), Ernestus and Warner propose a non-abstractionist ‘pure exemplar model’ that assumes ‘that all tokens of a word ever produced or perceived by the language user are stored in the mental lexicon with all their acoustic details’. They concluded that “wherever a theory is located on the theoretical continuum, it needs further specification of the production and comprehension mechanisms involved in the processing of reduced pronunciation variants. This specification cannot easily be obtained by means of research on the processing of just the full forms of words”. From a theoretical point of view, studies on the human processing of reduced speech can therefore make significant contribution.