EXISTENCE OF [V] AND [W] IN URDU LANGUAGE

NAUMAN KHALID

ABSTRACT

The primary purpose of this experiment was to investigate the existence of the sounds [v] and [w] in Urdu. Recordings of several sentences were made, and then analyzed through specialized software to construct the data.

Generally speaking, both sounds were found to be present in Urdu, and there percentage of existence was analyzed. The results were discussed in light of acoustic phonetics and earlier Urdu phonetic publications.

1. INTRODUCTION

1.1 Backgrounder

Urdu language belongs to the family of New Indo-Aryan (NIA) languages, which is a sub-branch of Indo-European languages. Urdu is spoken by at least 50 million people in more than ten countries as a first or a second language (the majority of speakers are in Pakistan and India). Urdu is similar to Hindi and both are derived from Khari-Boli or Dehlvi (Ethnologue.com; Hussain, 1997; Masica, 1993).

1.2 Problem Statement

As stated in the above section, Urdu is quite an old language (even though it was not called Urdu), and like any other language has gone through a number of semantic and phonological changes. These changes sometimes clarify the shady areas of the languages, and sometimes work against it.

Kachru explains that the phoneme [v] was originally part of Urdu (called Highly Persianised Urdu by him, and Khari Boli by some others (see above section)) (Kachru, 1987). At that stage Urdu was part of Old Indo-Aryan (OIA) languages (Masica, 1993, p. 99). During the transition of from OIA to NIA, the membership of [v] became questionable, and many linguists claimed of [w] being its replacement (Kachru, 1987; Masica, 1993). A confusion whether [v] or [w] is now present in Urdu was created and is present even now.

Although many publications e.g., Kachru, 1987; and Khan, 1997, have listed the sounds present in Urdu, but they have failed to scientifically classify whether [v], [w] or both exist in Urdu. Also left undetermined is the explanation of above consonants being allophones of each other or not. These sounds also seem familiar when spoken to a non-phonetician person. Their have been given many examples in the past indicating presence of either [v] or [w], but no data was ever given to prove their existence. In many instances [v] and [w] are used in allophonic distribution, and yet in others they are considered complimentary to each other.

This paper provides data and some analysis to determine the presence of [v] or [w] in Urdu language. Also it tries to address the distribution of [v] and [w].

2. LITERATURE REVIEW

The semi-vowel [w] is somewhat a shaky part of the NIA inventory. In a number of languages their occurrence is practically restricted to semi-predictable intervocalic glides (Masica, 1993). This suggests that [w] has further been changed over years to a more loose pronunciation that is somewhere in the middle of a vowel, a fricative and an approximant. Another explanation of above can be that the sound [v] never became a pure [w] and was left somewhere in the middle of the two.

The approximant [w] is strongest in the West amongst the NIA languages. The western [w] has a distinctive [v]-like allophone (although the contact is typically a loose one, between the upper teeth and the inside of the lower lip) before front vowels (Masica, 1993, p. 99-100). Kachru indicates that the approximant [w] is part of Urdu and the fricative [v] only exists in some highly Persianised Urdu accents (Kachru, 1987).

The fricative [v] is defined as a labio-dental fricative, pronounced by bringing the lower lip close to the upper teeth. Also the lip is shrunk inwards and the glottal pressure is increased. The air stream mechanism used is pulmonic egressive. The air on its way out passes between the teeth and the lower lip in a fricative manner (Khan, 1997, p. 115). The letter و orthographically represents this fricative.

The Urdu letter و (pronounced as 'vao') is also mapped onto the vowels [o], [ɔ], and [ʊ], therefore misleading some linguists to think that the English [w] approximant is also mapped onto by و, which is not correct. Word finally the [v] sound is pronounced lightly and is also sometimes changed to [o] (Khan, 1997, p. 115).

Many sounds have been borrowed from Old Indo-European and Old Indo-Aryan languages. The borrowed consonants roughly fall into three partially overlapping subsets. One of these sets is of those from Persian (in its Central Asian pronunciation): f, v, ʃ, z, χ, ɣ, q, are fairly well established in the languages of predominantly Muslim populations (though not equally in all) and the f, v, ʃ, z, of some others as well, preeminently common Hindi; all seven of the sounds mentioned are characteristics of Urdu (Masica, 1993, p. 105).

3. PROCEDURES

The problem required designing of an experiment, including recording and analysis of the data. For this purpose some subjects were selected for recording.

3.1 Subjects

The present experiment was performed on five adult male subjects, all native speakers of Urdu. For one subject five separate test recordings were made, thus giving a total of twenty sets of data. TABLE 1 lists the sentences given to all subjects for the experiment.

TABLE 1 The sentences recorded by the subjects for the experiment. Consult the Appendix for a detailed table IPA transcription (Table 5).

English Translation	Sentences in Urdu Script
My Purse was stolen.	میرا بٹوا گم ہوگیا۔
Hamid and Nasir are twin brothers.	حامد اور ناصر جڑواں بھائی ہیں۔
I said vaafir.	میں نے وافِر کہا۔
Strengthen yourself.	اپنے آپ میں قوّت پیدا کر۔
I said tarveej.	میں نے ترویج کہا۔

3.2 Preparation

To test whether the phones [v] and [w] exist in Urdu language, careful selection of words was made. These words were selected from known dictionaries (Farhang-e-Talaffuz, 1995; Feroz-ul-Lughat; Standard Twentieth Century Dictionary, 1992), and were analyzed so that normal native speakers were familiar with them. These words either contained [v] or [w] as described by the dictionaries. After selection of words, these were included in sentences listed in TABLE 1.

A computer system was setup that was able to record sounds through a high quality microphone, and an amplifier. A digital speech signal processing analysis software XWaves® by Entropic® was installed and tested. This software is able to provide information like spectrograms, power, and filtering to apply on a recorded speech signal. Another software Speech Analyzer® by Summer Institute of Linguistics® (SIL) was used for spectrogram analysis. This system was tested so that it gave minimal error on the data input, and was reliable through out the experiment.

3.3 Data Recording and Processing

The subjects were asked to repeat each sentence four times, in their natural style. The recordings were done at 22.0 KHz frequencies, and single channel. These recordings were saved as raw sampled data, in separate files.

The recordings were then processed to get spectrograms of the speech. The spectrogram can be generated by utility provided both in *XWaves®*, and *Speech Analyzer®*. In *XWaves®* the spectrogram can be viewed by opening the sampled data (.sd) file and selecting spectrogram from right-click menu on the waveform. In *Speech Analyzer®* the spectrogram can be opened by selecting the graph type of "Spectrogram A" or "Spectrogram B".

The phones [v] and [w] have distinct formant structure that makes them differ in spectrogram view. The spectrogram of phone [v] looks like the one in FIGURE 1. The recording to get this spectrogram was done in [ɑvɑ] context.

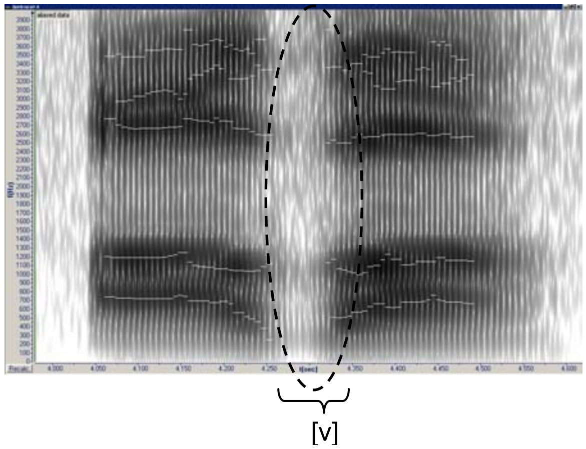


FIGURE 1 Spectrogram analysis of [ɑvɑ].

It is clear that here the formants F1 falls from 750 Hz to 550 Hz, before entering the phoneme /v/. On the other hand F2 remains stable at 1100 Hz. F3 is more or less at 2750 Hz. The phoneme /v/ has noise all over the spectrum, with voicing around 150 Hz. The noise starts around 2500 Hz, and increases in amplitude as the frequency increases. The formants have vanished in the time slot for phoneme [v].

The spectrogram of the phoneme [w] in the [ɑwɑ] context is shown in FIGURE 2. Here the formants can be seen during the production of [w]. Both F1 and F2 fall from 700 Hz and 1000 Hz to 400 Hz and 850 Hz respectively. The formant F3 has same frequency pattern of 2500 Hz, but at a lesser amplitude. There is no visible noise in the signal, and a huge gap between F2 and F3. F2 dips from both sides.

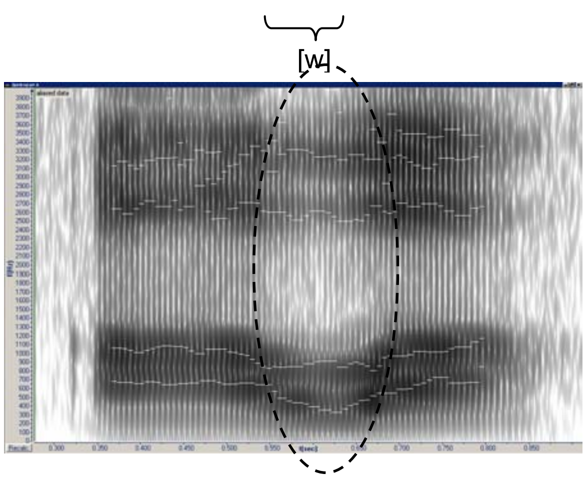


FIGURE 2 Spectrogram analysis of /ɑwɑ/, using Speech Analyzer Software.

In analysis of the data, the main cue for recognition of [v] and [w] was the spectrogram and formant transitions.

A secondary cue at recognizing whether it was [v] that was spoken is the spectrum. A spectrum was taken for both [v] and [w] for the interval that they were spoken. The intervals for [v] and [w] are also identified in FIGURE 1 and FIGURE 2 by brackets. Their spectrums are given in FIGURE 3. The higher peaks are visible in case of [w] (these are called F1, F2, and F3 respectively). In case of [v] no distinct high peaks can be seen after 500 Hz.

The values of formants for [w] from FIGURE 3 (b) are F1=309 Hz, F2=779 Hz, and F3=2722 Hz. Remember that more negative is lesser power. Please ignore the noise until 2000 Hz, as it was unavoidable during recordings. This noise is also present in all subsequent recordings.

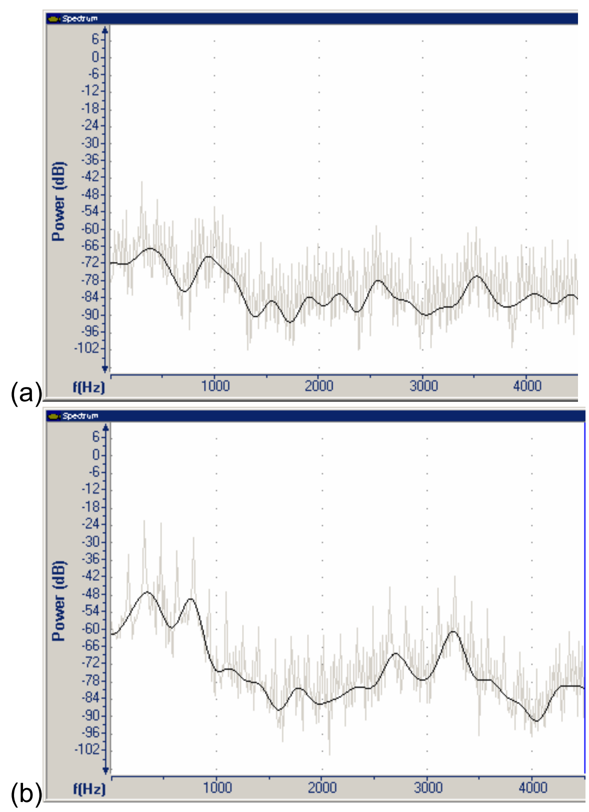


FIGURE 3 Spectrums of a) phoneme [v], and b) phoneme [w]. The amplitude at F1, F2 and F3 rise in (b), whereas in (a) the amplitude is more or less the same over all frequencies.

3.4 Experimental Conditions

All the recordings were done in minimal noise conditions, and frequencies lower than 70 Hz were ignored, to bypass any electrical noise introduced from the circuits. Even then some noise was introduced in the signal, which has degraded he quality of spectrograms and the spectrums. This noise is present until 2000 Hz of the signal. Since the major concern for conducting the experiment was the transition of formants, therefore the noise in signal is ignorable.

4. RESULTS

The formant values from FIGURE 1 and FIGURE 2 are given in TABLE 2. These values were taken as central values, and [v] and [w] were judged around these values.

4.1 The [w] approximant

The spectrograms of all the recordings were analyzed to study the formant transitions of entering into and out-of the [w] approximant. Also the spectrums of the approximant were analyzed to get the power (intensity) in decibels (dB) of each formant.

TABLE 2 Some Standard values of [w] and [v]. These are considered as central for other recordings.

	[w]	[v]
F1	309 Hz	Noise from 2000 Hz, to 4000 Hz. Voicing at around 120 Hz.
F2	779 Hz	
F3	2722 Hz	

4.2 The [v] fricative

The spectrogram of [v] shows noise especially visible above 25OO Hz. The formants cannot be determined in this case (because they do not exist!), so the only helpful measure is the noise level in the speech. This can also be determined by viewing the spectrum for the duration of the fricative.

4.3 The Transition Patterns

The TABLE 3 gives the average transition trends and noise distributions of [w] or [v] spoken by the subjects. These trends have been calculated from the spectrograms of the signal.

TABLE 3 Transition of formants or possible noise in spectrograms of [v] and/or [w]. The data given is averaged, and obtained from the experiment.

Word		Transition	
		IN	OUT
Batwa	F1	-	↗
	F2	-	→
	F3	-	→
Jurwan	F1	↘	↗
	F2	→	↗
	F3	↘	↗
Kuwat	F1	→	↗
	F2	→	↗
	F3	→	→
Tarveej	F1	↘	↗
	F2	↘	↗
	F3	→	→
Vafir	F1	-	↗
	F2	-	↗
	F3	-	→

4.4 The Formants and Intensities

The values of the formants and their intensity (power) are given in TABLE 4. These values are taken at the middle of phonemes [v] or [w]. In case of a pure [v] the formant values are missing. These values have been calculated from the spectrums of the signals.

The complete data extracted from the experiment of all the subjects is given in the appendix.

TABLE 4 Formants and intensity of the readings taken from the experiment.

Word			Averages
Batwa	Formants (Hz)	F1	456.2
		F2	1321.6
		F3	2204.8
	Power (dB)	1	-57.4
		2	-61.2
		3	-70.6
Jurwan	Formants (Hz)	F1	384.8
		F2	1205.4
		F3	2412.4
	Power (dB)	1	-49
		2	-63.8
		3	-67.2
Kuwat	Formants (Hz)	F1	347
		F2	807
		F3	2325.75
	Power (dB)	1	-51.25
		2	-56
		3	-73
Tarveej	Formants (Hz)	F1	284.8
		F2	1468.2
		F3	2264.8
	Power (dB)	1	-55.2
		2	-69.4
		3	-68.4

Vafir	Formants (Hz)	F1	375.2
		F2	1129.2
		F3	2261
	Power (dB)	1	-52.6
		2	-60.6
		3	-76
AVG.	***Formants (Hz)***	***F1***	***369.6***
		F2	***1186.2***
		F3	***2293.75***
	Power (dB)	***1***	***53.09***
		2	***62.2***
		3	***71.04***

5. DISCUSSIONS

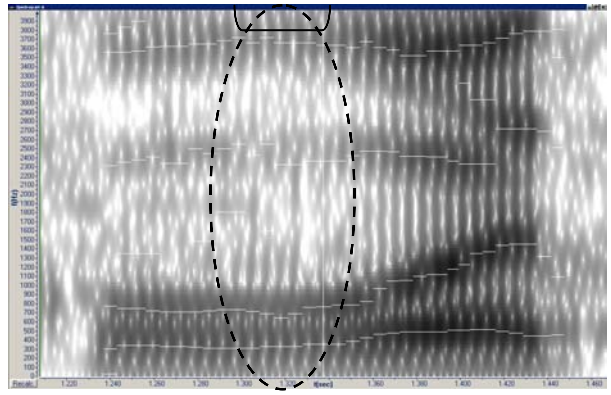
5.1 Result Summary

It was revealed in the experiment and its analysis that the native speakers of Urdu do not distinguish between the use of [v] or [w] in words. The message is conveyed equally well in both cases. Also enough data is present to suggest that being given freedom to pronounce anything from [v] to [w], the speakers generally tend towards a more loose pronunciation, somewhere between the approximant [w], the fricative [v], and the vowel [u].

5.2 Analysis

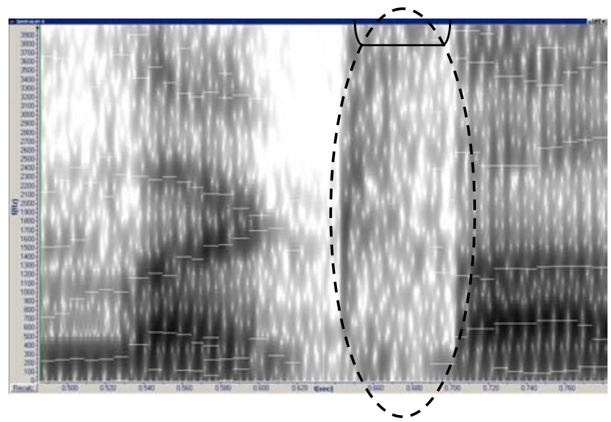
When the recordings were complete and the files properly saved, the words that had the "Problem Statement" in them (like Batwa, etc.) were selected to view the spectrograms. A sample of the spectrogram is given in FIGURE 4. The area that covers the [w] semi-vowel (approximant) is marked in the figure. The formants F1 and F2 decrease in frequency as they enter into the approximant. F3 on the other hand increases slightly. On exit the first three formants converge together. F1, F2 both increase, while F3 decreases. The values of formants can be located in the Appendix of this paper. This spectrogram shows that [w] was the sound uttered.

The above recording adds to the count of [w]'s uttered by the subjects. FIGURE 5 displays another spectrogram, this time of the word Batwa, taken from the recordings of the 5th subject. This particular subject has uttered a [v] instead of [w]. The voicing is visible at the bottom, and the noise above 2500 Hz (Please ignore any noise below 25500 Hz, as it was induced into the recordings from the electric circuitry. This noise is present in the closure of the previous stop).



[w]

FIGURE 4 Spectrogram of the word Kuwat, from the first subject.



[v]

FIGURE 5 Spectrogram of the words Batwa, from the recording of the 5th Subject.

Another important thing to note is the intensity level of noise, if [v] is present (to see if its evenly distributed). The spectrum of [v] in FIGURE 5 is showed in FIGURE 6.

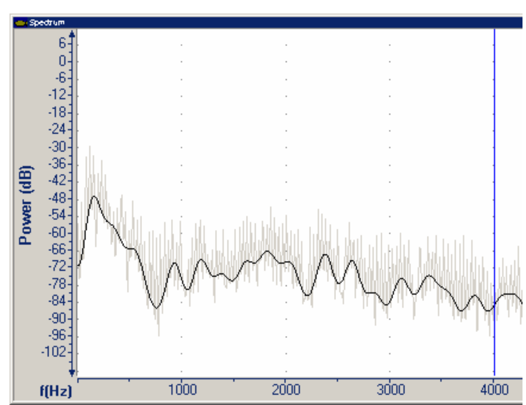


FIGURE 6 Spectrum of the [v] in Batwa uttered by Subject 5. Voicing below 200 Hz, and noise from 2500 Hz onwards (ignoring noise between 900 Hz and 2000 Hz).

All the recordings were analyzed and a concise result is available in TABLE 3, TABLE 4, FIGURE 8, and FIGURE 7.

5.3 Distribution Problem

Above discussion makes it possible to say that a non-empty set of *phones* exist in the range of [v] and [w] that map on to the phoneme /v/, but the set of phones do not have complimentary distribution (thus stripping them of the right to be called allophones). Another possible distribution could be that a non-empty set of *phonemes* exists that neither exhibits complimentary distribution nor parallel distribution (since no minimal pair of [v] and [w] can be found).

The above phenomenon is same as in the case of vowels. If a language has less number of vowels, then there place in the vowel quadrilateral is adjusted to keep symmetry and balance. Here since only one concept is present that is to be conveyed, the speakers have loosened their pronunciation sometimes approximating to the semi-vowel, and sometimes to the fricative.

The detailed results of the experiment, as given in the appendix, show that [v], and [w] were spoken by the subjects. FIGURE 8 and FIGURE 7 show the numbers graphically. The number of [w]'s spoken is much greater than [v], which makes one thing clear that native Urdu speakers do utter the sound [w] (FIGURE 7). Also in individual word counting (see FIGURE 8) the fricative is never in the lead.

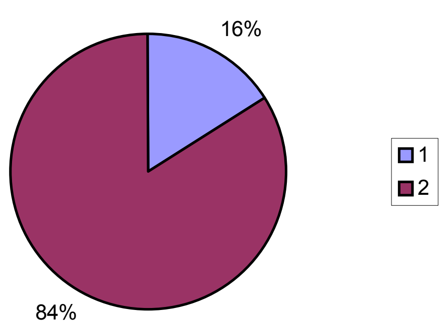


FIGURE 7 Overall percentage of the phones uttered. Legend 1: [v]; Legend 2: [w].

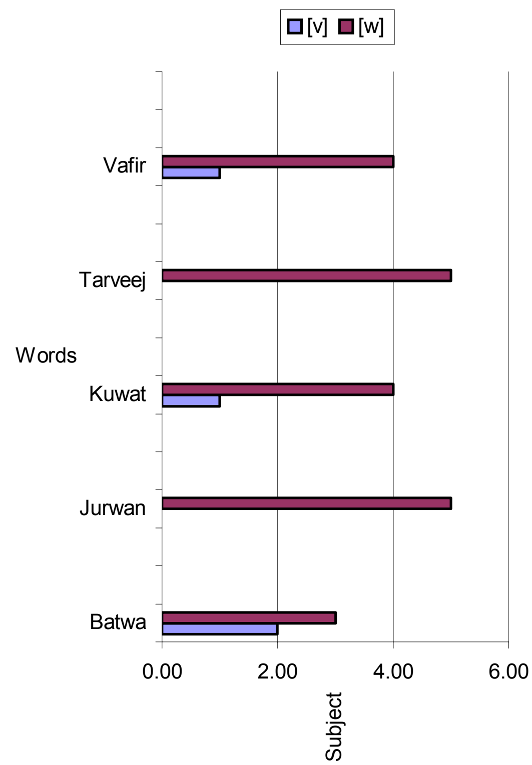


FIGURE 8 Word level pronunciation distribution. These numbers were gathered by successive comparison with formant values of TABLE 2.

The graph in FIGURE 8 is made from the data collected from the recording and the spectrograms. The data is given in Appendix in Table 2. The values were compared to those in TABLE 2, FIGURE 1, FIGURE 2 and FIGURE 3.

One of the subject— 5, was highly inclined to utter [v] and the percentage of [v] was greatly increased by his recording. If recordings of this subject are ignored from the experiment, then [v] utterance falls to 9%. This suggests that the distribution may be speaker dependant.

The distribution at phonological level that explains whether the utterance of [v] - [w] is speaker dependent or predictable according to any rule is not discussed in this paper. This aspect of the distribution of [v] and [w] should be investigated and published.

5.4 Transcription Problem

The dictionaries (Farhang-e-Talaffuz, 1995; Feroz-ul-Lughat; Standard Twentieth Century Dictionary, 1992) relate the transcription of و to the roman character "v". Also other references (Hussain, 1997, p. 152; Khan, 1997, p 115) refer to this sound as [v]. Since current experiment has provided ample proof that [w] exists in Urdu at least at the phonetic level, so these transcription standards have become questionable. Kachru and Masica on the other hand support the [w] transcription, as discussed in the Literature Review (Kachru, 1987, p. 59; Masica, 1993, p. p. 99).

Results of this experiment make it possible to say that the IPA symbol and sound [w] should be used for transcription purposes at the phonetic level, since majority of native speakers actually utter this sound.

6. REFERENCES

Bokhari, S. 1985. *Phonology of Urdu Language*. Royal Book Company, Karachi, Pakistan.

Bokhari, S. 1991. *Urdu Zuban Ka Sauti Nizam aur Tuqabli Mutalia.* Muqtadara Qaumi Zaban, Islamabad, Pakistan.

Ethnologue.com. Available at http://www.ethonologue.com/family_index.asp.

Farhang-e-Talaffuz. 1995. Muqtadara Qaumi Zaban, Islamabad, Pakistan.

Feroz-ul-Lughat. 1995. Feroz Sons (Pvt.) Limited, Lahore, Pakistan.

Hussain, Sarmad. 1997. Phonetic Correlates of Lexival Stress in Urdu.

Kachru, Y. 1987. "Hindi-Urdu," in *The Major Languages of South Asia, The Middle East and Africa*. Comrie, B (eds.). Routledge, London, UK.

Khan, M. 1997. *Urdu Ka Sauti Nizam*.