

# SYLLABLE TEMPLATES IN URDU LANGUAGE

**MUHAMMAD NOMAN NAZAR**

## ABSTRACT

Syllable templates play an important part in the syllabification algorithms and hence in the phonology of any language. The current paper investigates the syllable templates of Urdu. The results presented are based on a systematic analysis of a modest selection of words from a popular Urdu lexicon. The syllable templates extracted shed light on the influence of other languages over Urdu and raise some other interesting issues.

## 1. INTRODUCTION

In the phonology of any language, and indeed of the languages in general, the rules, constraints and requirements are best represented if they expressly refer to the syllable. According to the modern view, at least a rudimentary organization into syllables holds at all levels of the phonological representation. Hence, the gross shape of the maximal syllable is very much a part of the grammar of the language and has come to be known as the syllable template. The syllable templates are instrumental in the syllabification rules and algorithms of a language.

Very little work has been done on the phonology of the Urdu language to date. Hussain (Hussain, 1997) has done some work on the syllable templates of Urdu. The current study was aimed to investigate in detail the results presented therein and to extend them if possible.

## 2. LITERATURE REVIEW

### 2.1. Syllable

The role of the syllable in the phonological theory has been controversial (Kenstowicz, 1994, p.250; Trask, 1996, p. 345). It is a part of the conceptual baggage left from the traditional grammar. At various points in the history of linguistics, the syllable has been jettisoned in favor of a sparser theoretical vocabulary. Notably, during the early period

of the modern linguistics, influenced by the landmark work "Sound Patterns of English", the notion of syllable had no official recognition (Kenstowicz, 1994, p.250). However, through subsequent research, phonologists have come to appreciate that the syllable is an essential concept for understanding phonological structure.

The major reason for the controversiality of the syllable has been the difficulty of assigning it a proper definition. Despite the abundant evidence of the reality of syllables and thus clues about its nature, the syllable continues to resist the garb of a definition that neatly ties its phonetic and phonological character.

On the one hand, the native speakers find it easy to decide how many syllables are present in a given word or utterance; syllable-based writing systems have been in use for thousands of years and speech errors provide abundant evidence of the mental reality of syllables (Trask, 1996, p.345). While on the other hand, the lack of any uniform or direct phonetic correlates; variation of the exact shape of the syllable across languages and the effect of more superficial features of the language often obscuring the underlying organization of sounds into syllables (Kenstowicz, 1994, p.250), render it hard to define.

Historically, there have been various attempts to define the syllable phonetically. Trask lists a few including the chest-pulse theory: defining the syllable as a single respiratory movement; the prominence theory: defining the syllable as a single piece of prominence in the sound stream resulting from a combination of stress, pitch, length and intrinsic sonority; and also an attempt to define it as a single opening and closing of the vocal tract (Trask, 1996, p.345). However, none of these has proved adequate.

According to Trask, two approaches dominate the syllable in the modern phonology (Trask, 1996, p.345): -

- (1) Syllable is a unit of neural programming, which can be reconstructed by the hearer from a variety of clues, in spite of the absence of any single phonetic correlate.
- (2) The syllable is a purely phonological unit consisting of a single peak of intrinsic sonority (though with qualifications for cases like English spit, which has two peaks).

Summarily, the syllable can be thought of as an abstract unit of prosodic organization through which a language expresses much of its phonology, although the phonetic character of the syllable remains unclear (Kenstowicz, 1994, p.250). In much contemporary work, syllables are regarded as more fundamental than segments (Trask, 1996, p.345).

## 2.2. Syllable Structure

The syllable typically consists of a short sequence of segments, most typically, a single vowel or a diphthong, possibly preceded and/or followed by one or more consonants. It is now usual to divide the syllable into an onset and a rhyme, with the rhyme further subdivided into a nucleus (or peak) and a coda (Trask, 1996, p.345).

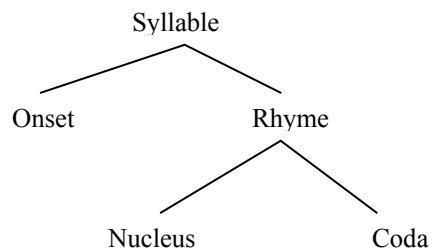


FIGURE 1 Structure of the syllable.

The nucleus plus coda form a tighter bound than the onset plus nucleus (Kenstowicz, 1994, p.252-253). Nucleus has the special status as the only obligatory constituent. It is the nucleus, which carries the tone or stress. Deletion of the nucleus typically relocates tone or stress, while loss of a consonantal onset or coda does not disturb the syllable

count and thus the location of tone or stress (Kenstowicz, 1994, p.253).

## 2.3. Syllabification

Syllabification, an analytical procedure for dividing a phonological representation into a well-defined sequence of syllables, is guided by two well-established rules:

1. *Maximal Onset Principle* states that a consonant, which may in principle occupy either rhyme or onset, will occupy onset position (Trask, 1996, p.217).

2. *Sonority Sequencing Principle* states that the sonority profile of the syllable must slope outwards from the nucleus. In other words, the level of sonority must rise as we proceed from the beginning of the syllable to the nucleus and fall as we proceed from the nucleus to the end in accordance with the sonority hierarchy (Trask, 1996, p. 328). (The terms 'sonority' and 'sonority hierarchy' are explained in the next section.)

These two rules guide the formation of complex onsets and codas. The notion of sonority perhaps deserves some elucidation. Sonority is a phonological property which lacks a simple phonetic correlate but can be associated with the output of the periodic acoustic energy of the segment or with the degree of aperture of the vocal tract (Trask, 1996, p.327). Phonologists agree that the entire class of speech sounds can be scaled in order of their intrinsic sonority as below, with the vowels as most sonorous and the stops as least.

*Sonority Hierarchy*, (from least sonorous to most sonorous)  
 oral stops, fricatives, nasals, liquids, glides, vowels

## 2.4. Syllable Templates

In the earliest works introducing the syllable into phonology, syllabification was performed by a battery of rules applied at a single point in the derivation or cyclically. Subsequently, an alternate view emerged according to which at least a rudimentary organization into syllables holds at all levels of the phonological representation. There is no point where syllabification originates or

ceases. There is ample phonological evidence that supports this view.

The explanation of the phenomenon of epenthesis, common in many languages, is a classic case. Epenthesis, or the insertion of a segment into a word where previously there was none, relocates the stress, and, in many cases, results in further phenomena like syncope, i.e., the deletion of a vowel. In the derivation of such cases, it was realized that syllabification is relevant at each step.

Ito (1986) has used epenthesis sites like these to argue for "templatic syllabification." According to this view, the gross shape of the maximal syllable is very much a part of the grammar of the language and has come to be known as the syllable template (Kenstowicz, 1994, p. 270).

A syllable template is usually formulated in terms of sequences of consonants and vowels. For example, Arabic language defines CV(V)C as its syllable template, where C denotes a consonant, V denotes a Vowel and the parentheses "( )" indicate optionality in the templatic elements. (Kenstowicz, 1994, p. 270).

Ito shows that templatic syllabification allows a simpler and more successful analysis of phenomena like epenthesis than the conventional syllabification-by-rule model. Syllabification consists in a directional mapping of the phonemic string to the appropriate positions in the template, matching vowels with V positions and consonants with C positions.

Syllable templates represent a fixed static set of constraints that dictate the syllable structure in the language concerned. Although, there is some evidence to suggest that templatic syllabification may need some revision (Kenstowicz, 1994, p. 276), the model is still relevant and valid for languages with an elementary syllable inventory like Arabic.

### 2.5. Syllable Templates of Urdu

Urdu belongs to the family of New-Indo-Aryan languages, which is a sub-branch of Indo-European languages. Urdu is similar to Hindi and both are derived from Khari Boli or

Dehlvi (Kachru, 1987). Although, both languages have the same origins and the same linguistic structure, Urdu phonetics and phonology have diverged from Hindi phonetics and phonology. The divergence is caused by the strong Perso-Arabic influence on Urdu and the strong Sanskrit influence on Hindi. (Hussain, 1997, p.40). According to Saxena, Urdu has retained its resilience in that it is characteristic of the language that it absorbs the words and phrases of other languages easily within its own grammatical structure. In fact it is the chief source of the continuing evolution of Urdu. On the other hand, Hindi has retracted back to its Sanskrit origins and has become more discriminating in allowing words of other languages to invade its boundaries (Saxena, 1978, p.2).

According to Saxena the languages that have been most influential in development of Urdu are Persian, Arabic, English, Portuguese and a host of local languages spoken in India at the time (Saxena, 1978, p.2-6). These languages continue to cast their influence in the direction of evolution of the Urdu language. It is therefore expected that an investigation of the syllable templates of Urdu should reveal results very similar to the templates of these languages. Also, the very prominent characteristic of Urdu, that of absorbing almost any word of other languages within its own grammatical framework suggests that the syllable templates of Urdu should comprise a rich reservoir. Moreover, it hints about the richness of its grammatical constructs also. Undertakings in the syntax, morphology and semantics of Urdu should therefore be worthwhile for the researchers. However, the current scope is restricted to the syllable templates only.

### 3. METHODOLOGY

The current investigation of relies on the "Jaibi Feroze-ul-Lughat, Feroze Sons, Lahore," which lists the Urdu words of common usage. There are more than twenty thousand of these words present in the dictionary. Five thousand of these words were chosen and examined to yield the syllable structure.

As in any language, there are some Urdu letters that are more popular or frequent than others in serving as the start of the word. The selection of the words aimed to provide equal representation to the words with the lower frequency and higher frequency. The number of words in the dictionary were listed down in descending order of starting-letter-frequency, indicated by the number of pages devoted to each alphabet. The words to be analyzed were then chosen alternately from the beginning and ending of this list. The reasoning behind this procedure is the assumption that if there are less number of words with a particular letter in the starting position than another, then there may be some constraint upon its formation that rejects most of the possibilities. Hence an investigation of the high frequency words alone may not elucidate this constraint. However, the high frequency words themselves represent a more generous constraint and therefore must not be completely ignored. An equal representation of both classes seemed the best choice and was adopted.

Another important consideration was the unbiased analysis of the structure of syllable. Since a native speaker of the language is best qualified to render judgement in this case, the analysis was based on the judgement of such a speaker, which in this case is the author himself. There is no standard transcription of Urdu words available, hence, care was necessary to avoid the bias that would have crept in had the author first transcribed the word and then attempted to analyze it. That is because the formal phonological training of the author would have come into play, in following the rules of syllabification and then making out syllables instead of judging the syllables from the spoken word and then figuring out their inherent templatic structure. Therefore, the words of Urdu from the dictionary were not transcribed, to obtain the best-unbiased data.

Another important aspect of the selection of words was to discriminate between the words of Urdu as opposed to those with a distinctly English origin. The reasons for this are given in the Discussion section.

## 4. RESULTS

The study reveals the following syllable templates of Urdu, listed in order of descending percentage frequency.

CVV	39
CVC	20
CVVC	15
CV	17
CVCC	3.2
VV	2.5
VVC	1.5
VC	1.0
CVVCC	0.3
V	0.4
VCC	0.05

As mentioned earlier, the syllable structure was based upon the judgement of a native speaker and the source upon which the investigation was based is "Feroz-ul-Lughat Jaibi(ref).

## 5. DISCUSSION

Urdu makes use of both short and long vowels, where the short vowels are denoted by diacritical marks and the long ones by special alphabets in script. It is therefore instructive to rearrange the given templates into two groups based upon the length of the vowel.

CVV	39	CV	17
CVVC	15	CVC	20
CVVCC	0.3	CVCC	3.2
VV	2.5	V	0.4
VVC	1.5	VC	1.0
		VCC	0.05

Each group is represented almost wholly by its first two members in terms of percentage frequency volume. Therefore it seems that a simple onset of one consonant, followed by a short or a long vowel, optionally followed by a simple coda of one consonant, i.e. CV(V)(C) is the makeup of the most natural syllable in Urdu.

Urdu language does not license a short vowel in word final position. That is, the (C)V template can never be at the end of a word.

Templates with the same popular makeup but having complex coda of two consonants, i.e. CV(V)(CC) are less preferred but comprise a significant fraction of the remaining percentage.

A syllable lacking an onset is least preferred and hence the template that comprises the lowest frequency is V(V)(C). The most striking feature of this makeup is the fact that the template VVCC containing a long vowel and a complex coda but lacking an onset consonant is prohibited completely.

The first and the most popular template CV(V)(C) is clearly the influence of Arabic on Urdu. From the knowledge of the author, Persian also possesses the same syllable template as being most popular. Although the syllable templates of Persian could not be located in a brief search of literature, it seems reasonable to assume that the CV(V)(C) template is the combined influence of Arabic and Persian on Urdu. The influence of Sanskrit in the formation of this syllable template could not be investigated due to ignorance of the author to the language concerned and non-availability of literature. The conclusion, however, is corroborated in part by Saxena (Saxena, 1978, p. 2-6), who maintains that the biggest influence that any language has had on the shaping of Urdu is the Persian language.

The evolution of a complex coda in the same makeup, i.e. the template CV(V)CC hints towards a Persian influence, though, this statement again remains to be corroborated.

There are limitations on the formation of the complex coda, primarily guided by the Sonority Sequencing Principle, which requires the sonority of the two consonants in the coda to be in descending order. However, these limitations do not seem as pronounced as mentioned in Hussain (Hussain, 1997, p.42), who maintains that "when there are two coda consonants, the first consonant in the coda is limited to a voiceless fricative or nasals and the second consonant is limited to stops." A common counter-example is the "rz" coda in many Urdu as in "Farz". As pointed out also by Hussain, detailed phonetic study of the

construction of complex codas still needs to be done.

Syllable weight should also play a significant role in the formation of Urdu templates. Apparently the bi-moraic syllables (i.e. syllables made of two units of weight) form the most preferred class of templates. Furthermore, the results also indicate a hint of extra-syllabicity (i.e. a unit consonantal or vowel, that has no moraic weight), as without it the Urdu would seem to possess a super-super-heavy syllable in the form of the CVVCC template. The moraic analysis of the Urdu syllables is therefore a promising direction of research.

The last category i.e. the template shorn of an onset cluster but otherwise following the same makeup V(V)(C)(C) again shows the same pattern. The V(V)(C) subset takes up most of the percentage volume of this category. While the introduction of another consonant in the coda, i.e. V(V)CC is constrained to such a limit that the one possibility VCC is severely restricted while the other VVCC is totally prohibited.

Perhaps the most remarkable result of these investigations is the fact that the complex onset mentioned by Hussain (Hussain, 1997, p.42) is totally absent in the results. However, a justification for this finding may be a bit controversial. The finding and its authenticity itself depend upon the discrimination practiced against words of the English origin imported in Urdu sentences. The reasoning for this is simple: the author judges himself not qualified in divining the structure of syllables in such words. This reasoning is based upon two important factors. Firstly, the author has been exposed to English language for close to fifteen years. Thus, the opinion of the author could never be unbiased in case of the Urdu-ized English words. Secondly, some pilot work done by the author to confirm or reject this bias, by putting the syllable structure of a select few words to the judgement of volunteers, has almost always resulted in disagreement. It seems that English words that have been in long usage in Urdu have had themselves resyllabified in most cases. The status of such Urdu-ized English words is therefore an open question. A detailed phonetic study needs to be done with

carefully chosen volunteers from amongst all education groups, as English education is the most important factor affecting the perceived syllable structure of such words. It is expected that the results of such a study will further enlarge the inventory of syllable templates of Urdu.

## 6. REFERENCES

Hussain, S., 1997, *Phonetic Correlates of Lexical Stress in Urdu*, Unpublished Ph.D. dissertation, Northwestern University, IL, USA.

*Jaibi, Feroz-ul-Lughat, Urdu*, Feroze Sons, Lahore.

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

Kenstowicz, Micheal., 1994, *Phonology In Generative Grammar*, Blackwell Publications.

Napoli, Donna Jo., 1996, *Linguistics, An Introduction*, Oxford University Press.

Saxena, Ram Babu., 1978, *Tareekh-e-Adab-e-Urdu*, Askari (trans.), Kashmiri(eds.).

Trask, R. L., 1996, *A Dictionary Of Phonetics And Phonology*, Routledge, London, UK.