FREQUENCY OF OCCURENCE OF PHONEMES IN KANNADA: A PRELIMINARY STUDY

[1]Sreedevi N., [2]Smitha K. Nair, & [3]Vikas M. D.

Abstract

Kannada is one of the traditional languages among the four Dravidian languages with a fine grammatical tradition. It is an enriched language with complex range of regional, social and stylistic variations. The frequency of occurrence of phoneme data is crucial to understand the language structure and also has wide applications in audiology and speech language pathology. The phoneme frequency is under research since 1930s. Studies on phonemes in various languages like English (Delattre, as cited in Edwards, 2003) Hindi (Ghatage, 1964), Oriya (Kelkar, 1994) were based on written source of materials. The purpose of the present study was to obtain the frequency of occurrence of various phonemes in Kannada using conversation samples. Participants included 21 fluent native speakers of Mysore dialect of Kannada in the age range of 20 to 50 years divided into 5 groups. Conversation sample of each of the five groups were recorded separately for 25-30 minutes. The samples obtained were transcribed using IPA transcription. Inter judge and intra judge reliability of phonetic transcription was evaluated for 10% of the recorded samples. Further it was analyzed using the SALT software for obtaining frequency of phonemes. Mean and standard deviation of frequency of phonemes of all the five samples were obtained. The results show that vowel /a/ was the most frequently occurring phoneme followed by /n/, /I/. /e/, /r/, /a:/, /d/, /l/, /u/, /g/ and /k/ in Kannada. Phonemes /h/, /s/, /p/, /tʃ /, /dʒ/, /ʃ/ were less frequent in conversational samples. Overall, vowels constituted 44.3% and consonants 55.3% of the data. The results obtained will aid audiologists and speech language pathologists in developing and updating the existing test material for evaluating various communication disorders and also for selection of treatment targets in such population. The study has implications in the area of linguistics and speech synthesis tasks also.

Key words*: SALT software, Dravidian languages.*

Phonology is the study of how the speech sounds are organized and the functions of those within the sound system of a particular language. According to Bloomfield (1914), phonology is the organization of sounds into patterns and phonemes are the minimal units of distinctive sound feature. In Webster's third new international (1961), a phoneme is defined as the smallest unit of speech distinguishing one unit from another, in all the variations it displays in the speech of one person or in one dialect as a result of modifying influences such as neighboring sounds or stress. Kruszewski was the first to use the term phoneme in late 1870s.

Most contemporary linguists view phoneme as the minimal bundle of relevant sound features. Trager and Smith (1951) proposed 45 phonemes for English such as /p/, /b/, /t/, /d/, /k/, /g/ etc. The phonemes are distinctive in all the languages but use of phonemes differs. English native speakers can clearly distinguish the phonemes of 'key' and 'caw' whereas the speaker of Hindi will be unlikely to hear any difference between the consonants of 'key' and 'caw', because his language does not force him to establish such a contrast. Kannada is one of the four major literary languages of the Dravidian family, the other three being Telugu, Tamil and Malayalam. It has a fine grammatical tradition and a very complex range of regional, social and stylistic variations: the Mysore/Bangalore dialect, the coastal dialect (Mangalore), the Dharwar dialect and Kalaburgi dialect (Upadhyaya, 1976). The Kannada lexicon has been enriched by uninhibited borrowing from several sources, majorly from Sanskrit, Hindi-Urdu, and English.

Determining frequency of occurrence of phonemes is foundation for linguistics and offers beneficial information to research and clinical fields. The frequency of occurrence of phonemes plays a crucial role in the development of linguistic theory in a number of areas including the grammatical relations, semantic structure etc. Audiologists use speech materials for hearing assessment and also in intervention. Several speech materials they use are in the form of word lists, involving phonetically balanced phonemes in a language. As these are language specific,

[1]Reader in Speech Sciences, All India Institute of Speech and Hearing (AIISH), Mysore-06, Email: srij_01@yahoo.co.in, [2] Clinical Linguist, AIISH, Mysore-06, Email: krishnasmi@gmail.com & [3]Research Officer, AIISH, Mysore-06, Email: vikasmdslp@gmail.com

they require most frequently occurring phonemes in that particular language (Egan, 1948; Campbell, 1965).

The frequency of occurrence of phonemes is in research since 1930s. As the phonological structure varies with the language use and the dialects, several studies were carried out in different languages. Delattre (as cited in Edwards, 2003) studied frequency of occurrence of phonemes in English. He considered 2000 syllables taken from different sources like narration and dramatization. Results indicated that most occurred vowels were /ə/, /I/, /æ/ and consonants were /t/, /n/, /r/ and /l/.

Up to 1970s the study of phonemes were mainly concentrating on the written context where the frequencies were mostly calculated from sources like newspapers, journals and scripts of the play. Mines, Hanson and Shoup (1978) used conversation in an interview and database of about 1,03,887 phoneme occurrences were obtained. The frequency of occurrences of phonemes were listed in descending order / a, n, t, i, s, r, l, d, ɛ /. These phonemes accounted for 47% of the total data.

There exist differences in frequency of occurrence of phonemes between written and spoken data. Hence Sandoval, Toledano, de la Torre, Garrote and Guirao (2008) studied and compared the syllabic and phonemic frequency in spoken and written context in Castilian Spanish. The results of the study indicated that /a/ and /e/ were the most occurring phonemes in both spoken and written context. Among consonants /s/ occurred about 8% and 7% in spoken and written context respectively. Followed by /s/ it was /n/ which occurred about 7% both in spoken and written contexts. /l/ and /d/ had similar frequency of occurrences. Sandoval et al., (2008) study indicated that even though there is no much difference in the frequency of occurrence of phonemes or syllables, there are slight variations when compared between two different modes i.e., written and spoken.

On similar lines, initial studies on Indian languages available were by Bhagwat (1961) where he calculated the phonemic and morphemic frequencies in Marathi. He considered this could be a source or database for devising a speed script. Ghatage (1964) calculated the phonemic and morphemic frequencies in Hindi using written source of materials. Results showed that vowels occurred more frequently than consonants. Followed by Hindi, Ghatage (1994) studied phonemic and

morphemic frequencies in Malayalam using 1,00,000 words from various written materials. The results indicated vowels /a/ and /I/ were the most occurring phonemes. Among consonants palatal nasal /n/ was most occurring followed by /k/ and /m/ respectively. Kelkar (1994) studied phonemic and morphemic frequencies in Oriya. The sources were similar to that of Ghatage's study. The results indicated that vowel /ə/ was the most occurring followed by /a/ and /I/. /r/, /k/ and /t/ were the most found consonant phonemes.

Frequency of occurrence of phonemes in Kannada was first studied by Ramakrishna (1962). He analyzed the speech sounds in Kannada in the written form. The results revealed that long vowels and aspirated phonemes are used relatively less frequently. Also reported that vowel /a/ is the highest occurring vowel and consonants like /r/, dentals /d/ and /t/ are the highly used consonants in Kannada.

Ranganatha (1982) considered 1,00,000 words from different written sources to calculate the frequency of occurrence of phonemes in Kannada. The results revealed that vowel /a/ was most frequently occurring followed by /I/, /u/, /n/, /d/, /a:/. The least frequently occurring phonemes were /pʰ/, /g/, /ɖʰ/, /kʰ/, /ʈʰ/. The data consisted of 48.6% of vowels and 51.4% of consonants.

In another study on Kannada, Jayaram (1985) carried out a longitudinal study for calculating the frequency of occurrence of phonemes considering written data from many sources like newspapers, journals etc. He listed out a series of vowels and consonants by rank order of their frequency of occurrence and the order was /a, I, n, r, u, d, ɔ, e, t, l/. Among consonants /n/, /r/, /d/ occurred most frequently. And the consonants constituted about 55% of the data.

India being a multicultural and multilinguistic background, there are dearth of studies about the frequency of phonemes in Indian languages. These studies would provide a database for developing speech materials for assessment and selecting treatment targets for various communication disorders, knowledge about the most frequently occurred phonemes can help in targeting those phonemes in therapy for hearing impaired. The phonetically balanced word lists that audiologists use for assessing auditory processing disorders like staggered spondaic words (SSW), for checking speech identification scores (SIS), speech in noise test (SPIN) and speech recognition scores (SRT) in routine audiological evaluations are based on such

phoneme frequency information and they are highly language specific. Such information can also help in the development of different aids and devices like text to speech converters for individuals with communication disorders.

The studies on frequency of occurrence of phonemes in Indian languages have considered only written source of materials including the earlier studies in Kannada (Ramakrishna, 1967; Ranganatha (1982); Jayaram, 1985). Also these results may not be appropriate at present as there are a lot of new words, modified and borrowed words used in day to day conversation. Also study by Sandoval et al (2008) indicated that there exist differences in phoneme frequency among spoken and written source of data. Hence the present study was planned to obtain the frequency of occurrence of phonemes in conversational speech samples in Kannada.

Method

Participants: Adult fluent native speakers of Kannada in the age range of 20 to 50 years with a minimum of 10 to 12 years of education in Kannada medium were selected for the study. All the participants were native speakers of Mysore dialect of Kannada and resided in Mysore city. Participants did not have any clinical history of speech, language, hearing or any neurological problems. A group of 3 to 5 participants were involved in each recording session. Table 1 shows participant details for each recording session.

Table 1: *Indicates number of participants involved in each recording session*

Recording Sessions	1	2	3	4	5
No. of Participants	4 (2 Males and 2 Females)	5 (4 Males and 1 Female)	4 (3 Males and 1 Female)	4 (1 Male and 3 Females)	4 (2 Males and 2 Females)

Instrumentation: A digital recorder (Olympus WS 100) was used for recording the conversation samples.

Procedure: The data was collected through conversation in controlled natural environments which lasted for about 25 to 30 minutes of duration. All the participants were made to sit comfortably facing each other and familiarize with the other participants of the session. The digital recorder was kept at equidistance from all the speakers. Participants were not given any particular topic about the conversation to avoid the repetition of words.

Any current topic of interest was initiated by the participants themselves and conversations were carried out for approximately half an hour. The participants were instructed to avoid words from other languages to the extent possible and to speak only in Kannada as naturally as possible. They were not restricted from using commonly used borrowed English words (E.g.: Bus, ticket, phone, car etc). A total of 5 sessions of general conversation were recorded for the study. Each recording session had different participants engaged in different topics of conversation. Table 2 shows the details of the topic of conversation in each recording session.

Table 2: *Indicates the topic of conversation in each recording session*

Recording Sessions	1	2	3	4	5
Topic of Conversation	Current Education system and political issues.	Family issues, corruption, present government	Casual talk on garments, provisions, postal dept and government officials.	Various cultural and therapeutic programs in Mysore city	Legal issues, family values and studies abroad

Data Analysis: The conversation samples were transcribed using IPA transcription. The repetitive words (eg: I may go to go to bank) and exclamatory remarks were not included in the analysis. English words that are very commonly used (eg. bus, car, bat, ticket etc) in day to day conversation were considered in the analysis.

The raw data obtained was analyzed using the software SALT RV version 9 for the frequency count. SALT (Systematic Analysis of Language Transcripts), is a computer based program designed to help in analyzing and interpreting language samples from one or more speakers during a communicative interaction. It can be used to analyze samples from everyday speech like conversation and narration. Using SALT program, clinicians and researchers can analyze lexical, syntactic, semantic, pragmatic, rate and fluency categories and transcribe samples in a common format. This software analyses the transcribed sample for different parameters such as MLU (Mean Length of Utterance), NDW

(Number of Different Words), TTR (Type Token Ratio) etc. It also gives information about the frequency of words, morphemes, grammatical categories, etc.

Using SALT, an individual's language sample may also be compared to a reference database of language measures. It helps in managing the steps of word count, phoneme count, which can be preloaded with the editable database. So a database of Kannada phonemes was prepared and saved in the editable standard wordlists of SALT software. The database consisted of all the phonemes available in Kannada adapted from Sridhar (1990) which was then modified according to SALT conventions. The whole conversation data is edited with spacing after every phoneme and this edited file is loaded into SALT. The Salt software compares the loaded phoneme file with the database and provides the phoneme count.

Inter-judge and Intra Judge Reliability:
Sample of each conversation recording was subjected to inter judge and intra judge reliability measures. Three judges including two post graduate speech language pathologists and a clinical linguist served as judges for determining inter judge reliability measures. Judges were instructed about the phonetic coding procedure and the material to be transcribed before the actual transcription procedure. For inter judge reliability 10% sample of the 30 minutes recording of each sample was transcribed by each of the three judges. The recorded samples were played to the judges independently. They were not allowed to discuss about the transcription of the sample before or after the task. For intra judge reliability 10% of each of the 5 recordings were transcribed and analyzed by one of the authors (a speech language pathologist) after transcribing all the samples completely once. As each individual participant was given equal opportunity in the conversation recording, inter and intra judge reliability was measured for each participant rather than for each recording. The statistical procedure, Cronbach alpha test was used and a reliability index (alpha) of 0.87 was obtained for inter judge and 0.94 was obtained for intra judge reliability.

<div align="center">

Results

</div>

The present study aimed to obtain the frequency of occurrence of phonemes in conversational speech in Mysore dialect of Kannada. The results of phoneme count obtained from SALT software were tabulated.

A total of 69,624 phoneme counts accounted the data collected from conversation samples of five groups of participants. Figure 1 represents the number of phonemes obtained from the each of the five recording sessions. Out of the five recordings, recording 1 (R1) had the maximum number of phonemes and Recording 2 (R2) had the minimum number of phonemes.
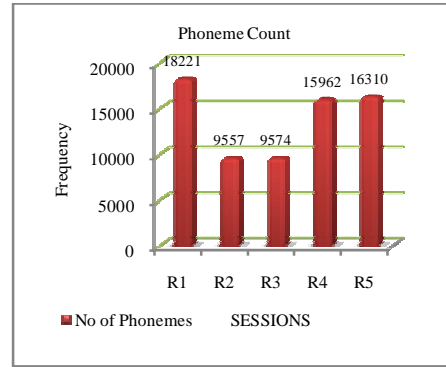

Figure 1: *Shows the total number of phonemes obtained in five conversation samples. R1-R5 indicates five recording sessions.*

Using descriptive statistics, the mean and standard deviations of the frequency of occurrence of vowels and consonants were calculated. Figure 2 represents the mean frequency of vowels and consonants calculated from all the 5 samples. The total vowel count was 30,946 which constituted 44.3% of the total data whereas consonants accounted for 38,618 which is 55.3% of the total data.
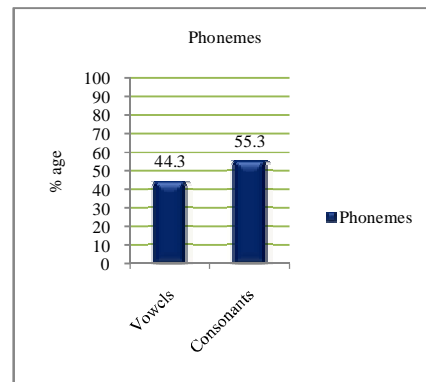

Figure 2: *Mean frequency of occurrence of vowels and consonants in percentage.*

Table 3 shows frequency of occurrence of the most frequently occurring phonemes along with their percentage of occurrence in each of the five recordings. From recording session 1 (R1) total number of 3671 words were obtained which consisted of 18,221 phonemes. The most frequently occurred phoneme was /a/ (2819) followed by consonant /n/ (1563). The frequency

of occurrence of other phonemes in decreasing order are /I/, /r/, /e/, /d/, /t/, /l/, /u/, /k/ and /g/. From recording session 2 (R2) 1860 words were obtained which consisted of 9,557 phonemes. The frequency of occurrence followed a similar trend as of R1 as vowel /a/ (1313) was followed by consonant /n/ (710). The frequency of occurrence of other phonemes in decreasing order are: /a/, /n/, /I/, /e/, /r/, /a:/, /d/, /l/, /u/, /t/, /k/ and /g/.

A total number of 2,153 words were obtained from recording session 3 (R3) which consisted of 9,574 phonemes. The frequency count obtained were similar to first two recording sessions with /a/ (1289) being most frequently occurred phoneme followed by /n/ (765). The frequency of occurrence of other phonemes in decreasing order are: /I/, /e/, /d/, /r/, /a:/, /l/, /t/, /u/, /k/ and /g/.

Table 3: *Shows the frequency and percentage (in parenthesis) of occurrence of the most frequently occurring phonemes for each recording session*

	/a/	/n/	/I/	/e/	/r/	/a:/	/d/	/l/	/t/	/u/
R 1	2819	1563	1246	1090	1094	1065	1013	847	889	767
	(15.4)	(8.5)	(6.8)	(5.9)	(6)	(5.8)	(5.5)	(4.6)	(4.8)	(4.2)
R 2	1313	710	688	592	542	502	490	522	352	381
	(13.7)	(7.4)	(7.2)	(6.1)	(5.6)	(5.2)	(5.1)	(5.4)	(3.6)	(3.9)
R 3	1289	765	677	582	527	500	527	478	431	395
	(13.4)	(7.9)	(7)	(6)	(5.5)	(5.2)	(5.5)	(4.9)	(4.5)	(4.1)
R 4	2196	1214	1147	909	813	823	882	929	806	717
	(13.7)	(7.6)	(7.1)	(5.6)	(5)	(5.1)	(5.5)	(5.8)	(5)	(4.4)
R 5	2264	1264	1158	936	794	863	836	778	687	883
	(13.8)	(7.7)	(7.1)	(5.7)	(4.8)	(5.2)	(5.1)	(4.7)	(4.2)	(5.4)

A total number of 1860 words were obtained from the recording session 4 (R4) which included 9,557 phonemes. The trend remained same with /a/ (2196) being most frequently occurred phoneme followed by /n/ (1214). The frequency of occurrence of other phonemes in decreasing order are: /I/, /l/, /e/, /d/, /a:/, /r/, /t/, /u/, /g/ and /k/.

A total number of 3671 words were obtained from recording session 5 (R5) which consisted of 18,221 phonemes. Same as in the previous four recordings, the most frequently occurred phoneme was /a/ (2264) followed by /n/ (1260). Other frequently occurring phonemes in decreasing order are: /I/, /e/, /u/, /a:/, /d/, /r/, /l/, /t/, /g/ and /k/.

Table 4 provides the mean percentage and standard deviation of all the phonemes from the 5 conversation samples. The standard deviation of the data was minimal, across the sessions and hence the reliability of frequency of phonemes is high.

Table 4: *Shows mean percentage and standard deviation of vowels and consonants*

VOWELS		CONSONANTS			
	Mean % (SD)		Mean % (SD)		Mean % (SD)
/a/	14.06 (0.8)	/n/	7.87 (0.4)	/s/	1.72 (0.2)
/I/	7.08 (0.14)	/r/	5.43 (0.4)	/ʈ/	1.61 (0.38)
/e/	5.94 (0.21)	/l/	5.14 (0.4)	/ḻ/	1.29 (0.28)
/a:/	5.35 (0.27)	/t/	4.47 (0.5)	/j/	1.21 (0.14)
/o/	2.08 (0.24)	/g/	3.46 (0.1)	/p/	0.94 (0.25)
/e:/	1.83 (0.73)	/k/	3.34 (0.1)	/ʃ/	0.61 (0.31)
/o:/	1.26 (0.21)	/m/	2.81 (0.2)	/dʒ/	0.44 (0.12)
/i:/	0.95 (0.27)	/v/	2.59 (0.1)	/tʃ/	0.36 (0.11)
/u:/	0.57 (0.1)	/ɖ/	2.41 (0.3)	/ɳ/	0.27 (0.05)
/ə/	0.56 (0.88)	/b/	2 (0.2)	/kʰ/, /ʈʰ/	0.03 (0.02)
		/h/	1.87 (0.4)	/ŋ/, /pʰ/, /ɖʰ/	0

Figure 3 represents the mean percentage of the most frequently occurring twelve phonemes. On overall observation the frequency of occurrence of phonemes in decreasing order are: /a/, /n/, /I/, /e/, /r/, /a:/, /d/, /l/, /t/, /u/, /g/ and /k/. The consonants /m/ and /v/ occurred for 2.7% of the total data. Vowel /o/ occurred 2% whereas phonemes /h/, /s/, /j/ occurred less than 2% and /p/, /tʃ/, /ʃ/, /i:/, /o:/ occurred less than 1% each of the total data. Consonants /ŋ/, /pʰ/, /ɖʰ/ did not occur in the five recorded conversational samples though they are present in the Kannada phoneme system. The aspirated phonemes were amply seen. Diphthongs /ai/ and /au/ occurred for less than 1% of the total data. /ai/ occurred for 0.2% and /au/ occurred 0.06% of the total data.
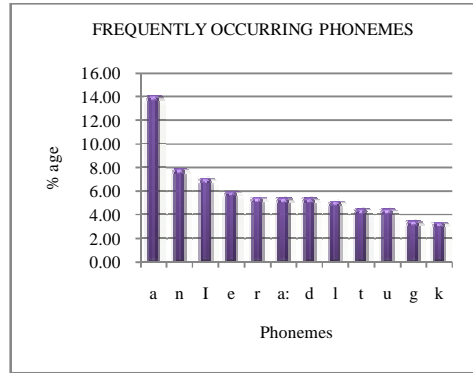
Figure 3: *Mean percentage of frequency of occurrence of frequently occurring phonemes.*

Discussion

The results of the present study had several salient findings. The results indicated that the Kannada conversational speech consists 55.3% of consonants and 44.3% of vowels. This suggests that the consonants are used more than vowels in our day to day conversation. This is in consonance with the earlier reports of several languages. Yegerlehner and Voegelin (1957) observed similar findings in five out of nine languages. The results of Delattre's (as cited in Edwards, 2003) study in English and Jayaram (1985) in Kannada also showed that the consonants are used more than vowels.

In the present study, the most frequently occurring phonemes in descending order are /a/, /n/, /I/, /l/, /e/, /d/, /a:/, /r/, /t/, /u/, /g/ and /k/. These phonemes constituted 70.2% of the total data where as in English it was 47% of the total data (Mines, Hanson & Shoup, 1978). The order of the frequently occurring phonemes in the present study is in contrary to the reports of Ramakrishna (1967) in written Kannada. Though he found that vowel /a/ occurred most frequently, the predominantly occurring consonants were /r/, /d/ and /t/.

In the present study the frequent consonants /n/, /l/, /r/, /t/, /g/ and /k/ constitute 31% of the 55% total consonantal data. /n/ was found to be the most occurring consonant which constituted 7.9% of the total data. The findings of Jayaram (1985) in written Kannada also showed /n/ as the most frequently occuring consonant but was followed by /r/, /d/, /t/ and /l/ indicating a difference in the sequence of the frequently occurring phonemes across these two studies. In English, the results of Sandoval et al (2008) showed that /s/ was the most frequently occurred consonant whereas Delattre reported it to be /t/ as the mostly frequently occurring consonant. The results of these studies clearly indicate that the frequency of use of consonants varies across languages and also across modes that is written and spoken context.

The aspirated consonants were rarely seen in the present study. Aspirated /pʰ/, /ɖʰ/, /kʰ/, /ʈʰ/ occurred for 0-0.05% of the total data. This is a significant finding of the present study as the other studies by Ramakrishna (1968), Ranganatha (1982), Jayaram (1985) showed the occurrence (0.5 % to 1.35%) of these aspirated consonants in written context. This indicates that although the aspiration is present in Kannada phoneme system, it is sparingly used in conversation.

Considering the vowels in the present study, vowel /a/ was the most frequently occurring vowel followed by /I/, /e/, /a:/ and /u/. A similar trend was observed by Mines, Hanson and Shoup (1978) in English. However, the order was relatively different in the study by Ranganatha (1982) in written Kannada which showed that /a/ was the most frequently occurring vowel followed by /I/, /u/, /a:/ and /e/.

The present study yielded similar order of phonemes in all the five recordings, i.e. vowel /a/ was the most frequently occurring phoneme followed by nasal /n/ and the relatively less occurring consonants were velars /k/ and /g/. This indicates that there was a consistency in the occurrence of phonemes in spite of the change in the topic of conversations across the different recording sessions. However this needs to be demonstrated on a larger sample of conversation recordings.

Conclusions

The frequency of occurrence of different phonemes in Kannada was determined using conversational samples and the most frequently occurring phonemes in the descending order are /a/, /n/, /I/, /l/, /e/, /d/, /a:/, /r/, /t/, /u/, /g/ and /k/ whch constituted 70.2% of the total data. The less frequently occurring phonemes were /m/, /v/, /o/, s/, /p/, /h/, /tʃ/, /dʒ/, /ʃ/, /i:/ and /o:/. The information obtained in this study is applicable not only in the diagnostic and therapeutic aspects of communication disorders but also in the area of linguistics and speech synthesis which are a few to mention.

References

Bhagwat, S. V. (1961). Phonemic frequencies in Marathi and their relation to devising a speed-script. Pune: Deccan College

Bloomfield, L. (1914). *Introduction to the study of Language*. New York: Henry Holt.

Campbell, R. A. (1965). Discrimination and word test difficulty. *Journal of Speech and Hearing Research, 8,* 13–22.

Edwards, H. T. (2003). *Applied Phonetics: The sounds of American English (3rd edition).* Canada: Delmar Learning.

Egan, J. P. (1948). Articulation testing methods. *Laryngoscope, 58,* 955–991.

Ghatage, A. M. (1964). *Phonemic and Morphemic frequencies in Hindi.* Poona: Deccan College Postgraduate and Research Institute.

Ghatage, A. M. (1994). *Phonemic and morphemic frequencies in Malayalam.* Mysore: Central Institute of Indian Languages.

Jayaram, M. (1985). Sound and Syllable distribution in written Kannada and their application to Speech and Hearing. *Journal of All India Institute of Speech and Hearing, 16,* 19-30.

Kelkar, A. R. (1994). *Phonemic and Morphophonemic frequency in Oriya.* Mysore: CIIL.

Mines, M. A., Hanson, B. F., & Shoup, J. E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech, 21,* 221-41.

Ramakrishna, B. S. (1962). *Some aspects of the relative efficiencies of Indian languages: a study from information theory point of view.* Bangalore: Indian Institute of Science.

Ranganatha, M. R. (1982). *Morphophonemic analysis of the Kannada language.* Mysore: CIIL.

Sandoval, A. M., Toledano, D. T., de la Torre, R., Garrote, M., & Guirao, J. M. (2008). Developing a Phonemic and Syllabic Frequency Inventory for spontaneous spoken Castalian Spanish and their comparison to Text-based Inventories. *Proceedings of the Language Resources and Evaluation Conference.* Available at: www.lrec-conf.org/proceedings/lrec2008/pdf/283_paper.pdf (accessed: 14 March 2012).

Sridhar, S. N. (2007). *Modern Kannada Grammar.* New Delhi: Manohar.

Trager, G. L., & Smith, H. L. (1951). *An outline of English structure* (Studies in Linguistics, occasional papers, No. 3). Norman, OK: Battenberg Press.

Upadhyaya, U. P. (1976). *A Comparative Study of Kannada Dialects: Bellary, Gulbarga, Kumar, and Nanjangud Dialects.* Mysore: Prasaranga, University of Mysore.

Webster, N., & Gove, P. B. (1961). *Webster's third new International dictionary of the English language.* London: G. Bell.