

Inter-Lingual Homophone Retrieval in Bilinguals

Ramya Maitreyee¹ and S.P. Goswami¹

Abstract

There has been a debate since decades as to whether bilinguals are like two monolinguals within the same individual or they share a same lexicon. Is the representation in the brain shared or are they completely separate is still being researched upon. Inter-lingual homophones which are words with similar form but different meanings across languages provide a real challenge to a bilingual's recognition system. The present study was aimed to investigate the inter-lingual homophone retrieval abilities in normal bilinguals. Data were collected from a total of 46 participants who were bilinguals. Twenty paired words which consisted of inter-lingual homophones were presented. The participants were asked to listen carefully to the pairs of words which were in two different languages that they knew and were asked to write the meaning of each word. The results revealed that the native language was more dominant for retrieval of the meanings of the words for children (10-12 years) and adults (35-45 years) in both Kannada and Hindi native speakers. However, teenagers (17-19 years), performed similarly in both the languages. It is hypothesized that a bilingual has separate lexicons for L1 and L2 in the younger age, and then an interaction occurs between both the lexicons and finally the most used language becomes dominant. The results are also discussed in terms of the Revised Hierarchical Model (Kroll & Stewart, 1994).

Keywords: Bilinguals, Interlingual-homophone, Lexicon, Retrieval.

/mara/, /mara/..... Are they from two different languages? How then is multiple language structure organized in relation to one another in the brain? How do we then acquire, organize and process lexical knowledge when more than one language is involved? Whether the two languages of a bilingual child to which s/he is exposed since birth are mixed in his mind and the age at which separation of the two languages occurs is one of the current issues in early bilingualism. Woutersen (1996, 1997) delineates bilingualism as sub-ordinative, compound and coordinate which are in part associable with different stages or levels of bilingual development. Woutersen (1996) conducted lexical decision experiments involving repetition priming with three groups of Dutch learners of English of different levels of proficiency in English (secondary-school level, university level, and near-native level). The participants were asked to make a lexical decision regarding the phonetic string (which was either in Dutch or English), which they heard over headphones if it was or was not a real word. The results obtained showed an inter-lingual priming effect among the intermediate (secondary-school) learners only for cognates and not for non-cognates, whereas

among the high-proficiency and near native learners there was an inter-lingual priming effect for both. In case of less-proficient learners, only when there was an inter-lingual overlap at the formal level did priming occur, whereas for the more proficient subjects, overlap at the semantic level alone triggered priming. Woutersen (1996, 1997) concluded that the less proficient learners exhibit bilingualism of the sub-ordinative type, whereas the highly proficient and near-native learners exhibit bilingualism of the compound type, concluding that an individual's lexical organization moves from the former category to the latter as proficiency increases.

Contradictory evidence exists concerning the mode by which a fluent bilingual stores and retrieves words in each of the languages. Curtis (1978) suggested that bilingual's storage is shared but organized and retrieved primarily through his first language (L1). The acquisition and processing of form and meaning of second language (L2) lexis have been a prominent theme of debate. One school of thought says that the basis of the operations of the L2 lexicon is phonological rather than semantic, that "while in the native speakers mental lexicon there are strong semantic links between the words, the

¹ All India Institute of Speech and Hearing, Mysore, India

connection between words in additional languages are primarily phonological" (Fromkin, 1971; Hatch, 1983; Laufer, 1989; Sondek, 1982).

A review of literature related to language disorders show that most individuals lose or recover multiple languages equally (Paradis, 1987), but some recover one before the other, and some recover either L1 or L2. These results state that two or more languages may have different representation or levels of activation.

Inter-lingual homophones are terms that sound the same or very similar but do not have the same meaning in both languages. One of the most critical issues in lexical research involves the representation of homophones, homographs, cognates and false cognates. According to Bybee (1985) and Kirsner (1986), recognition of false cognates involves representations in separate lexicons. In a study done by Lalor and Kirsner (2001), interpretation of false cognates was constrained by meaning rather than language and the results imply that lexical representation in bilinguals is organized along morphological lines and is not governed by language when processing an inter-lingual homograph (such as French/English coin). Bilinguals access its meaning in both their languages when taking vocabulary tests in one of their languages. In favor of L1-L2 lexical integration, Kirsner et. al, (1984) found that the presentation of a stimulus in one of bilingual subjects languages primes his/her response to a corresponding stimulus in his/her other language. Green (1986) found that when a person has a reasonable command of two languages, lexical items are subconsciously activated in both the languages, those in the

language not required being suppressed. The usage of only a single language at a time by a bilingual indicates the separation of the respective lexicons. In studies of language loss due to brain damage in bilinguals/multilinguals also supports the representation of different lexicons. In such instances of language loss, often all languages known to the individuals are effaced and are then recovered one by one (the L1 not necessarily appearing first). Grainger (1994) cites that content plays a role in accessing words in one language or another, and those bilinguals have more difficulty in processing mixed sequences of words than sentences presented only in a single language. This finding reflects that the opening word of the sentence 'switches on' the lexicon in either of the bilingual's languages, and that all lexical searching takes place initially in that particular lexicon.

The identification of inter-lingual homographs, within lists of words from one or other of a bilingual's languages depends on the frequency of that form in that particular language (Gerard & Scarborough, 1989). An inter-lingual homograph activates target words in both of the bilingual's language in lexical-decision tasks. Beauvillian and Grainger (1987) presented pairs of isolated words to participants and measured the reaction time required to identify the second of the pair in same and different language conditions and found that the processing of a French word is slower when preceded by an English word except where the words in question have distinctively French or English orthographic features. This indicates that contextual information about

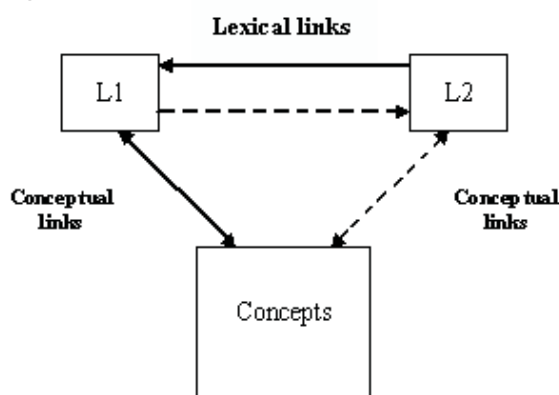


Fig. 1. Revised hierarchical model (Source: Adapted from Kroll & Stewart, 1994)

language choice is used on a word-by-word basis rather than at a more general level. The inhibiting effects of changing language was much stronger when the first word was presented in the bilingual's dominant language than when it was in his/her weaker language which appears to indicate a basic level of activation for words in the dominant language which is higher than that for words in the weaker language.

Grainger (1994) found that during an initial phase of recognition, both the languages are activated irrespective of the language context and then contextual information comes into play in order to reduce levels of activation of lexical representation in the inappropriate and/or to raise levels of activation in the appropriate language. The effects of contextual information are greater when derived from dominant language content. However, he does not support the idea of bilingual lexical search as proceeding sequentially, language by language; rather supports the activation of all relevant lexical representations in both languages in response to a given stimulus, and a consequent selection is based on the contextual information and the distinctiveness of the stimulus. According to him each language is separately represented, the lexicon of each language being more or less activated by the outcomes of lexical search and according to the degree of strength of each language. Gerard and Scarborough (1989) have quoted that a lexicon-specific access exists (i.e. access to one language at a time) in a bilingual's mental lexicon in contrast to some authors who have suggested a non-selective access (i.e. access of both the languages simultaneously). Lexical access can be language specific and is achieved by considering only the activation level of the lexical nodes (Costa, 1999).

A number of models have also been proposed to explain the phenomenon of visual orthographic processing in bilingual individuals and to study whether the bilingual has separate or shared representation. Kroll and Stewart (1994) proposed the revised hierarchical model (Figure1) to capture the implications of early reliance on L1 for the form of word-to-concept connections. The model merges the word association and concept mediation alternatives

into a single model in which the strength of the connections between words in L1 and L2 and concepts is proposed to take on different values. The initial dependence on L1 to mediate access to meaning for L2 words is assumed to create strong lexical level connections from L2 to L1. However, at a lexical level, the connections from L1 to L2 are not assumed to be particularly strong because there is a little need for the learner to use L2 in this way. Likewise, the model assumes that connections between words and concepts are stronger for L1 than for L2. A number of empirical findings support the predictions of the revised hierarchical model. More critically, only translation from L1 to L2 was influenced by the presence of semantic information. The absence of semantic effects in the L2 to L1 direction of translation suggests that it was possible for bilinguals to translate directly at a lexical level.

All the models proposed till date have mostly examined the representation of both the languages in a bilinguals brain using homographs, cognates, etc. and hence have evaluated the visual orthographic processing in them. Research in alphabetic languages (E.g. Dutch, English) has frequently indicated a non-selective lexical access of the orthographic mental lexicon in bilinguals. Indian milieu is a very complex one as it is a multilingual, multi-ethnic, and multi-religious country. There is a lack of understanding of the complex but structured network of linguistic communication and bilingualism in India. Most of the Indian languages are purported to employ a syllabic or semi-syllabic writing system where a unique shape is used for each syllable. It has also been noted that a good amount of phoneme grapheme correspondence occurs in Indian languages than in alphabetic languages. However, whether the bilingual processes the languages simultaneously or sequentially when presented with inter-lingual homophones need to be researched upon, which provides the motivational grounds to carry out the present study. The present study was aimed to investigate the language of dominance and its pattern in bilinguals using inter-lingual homophones i.e. words that have a similar pronunciation but different meaning within or across languages.

Method

Participants

A total of 46 participants (21 males and 25 females) participated in the present study. Participants were explained the purpose and procedures of the study, and an informed verbal consent was taken. The educational level of all the participants ranged from 5th grade to graduate level. The study group consisted of both native speakers of Hindi (22 participants) and native speakers of Kannada (24 participants). The native participants of Hindi were born and brought up in the state of Karnataka and were fluent in communicating in both the languages. Hindi, an Indo-Aryan language and the national language of India, is most commonly spoken in the northern parts of India. Kannada is a Dravidian language spoken mostly in the southern part of India, i.e., in the State of Karnataka. The participants were divided into three age groups. Group I included participants in the age range of 10 to 12 years with a mean age of 10.8. Participants of group II were in the age range of 17 to 19 years with a mean age of 18.5 years. Group III participants were in the age range of 35 to 45 years with a mean age of 40 years.

Although all the participants were proficient and comfortable in using both the languages they used their native language (L1) for day-to-day communication and their second language (L2) for academic purposes. The Australian Second Language Proficiency Rating (ASLPR) was administered and the participants having vocational level of proficiency were selected for the study. The criteria for selection of subjects were as follows:

1) Group I: All the participants had acquired Kannada/Hindi as their native language and had studied Hindi/Kannada respectively at least for a minimum of one year for academic purposes.

2) Group II: All the participants were in their pre-university level and had acquired Kannada/Hindi as their native language and had studied Hindi/Kannada respectively for a minimum of 8 years.

3) Group III: All the participants had a minimum educational qualification of Bachelor's degree and had acquired Kannada/Hindi as their native

language and were exposed to Hindi/Kannada respectively for at least a minimum of 8 years.

It was ensured that there was no past/present history of any neurological or psychological problems in participants of all the groups. Also there were no obvious sensory deficits in any of the participants.

Materials

1. *Stimuli used:* A non-standardized list of 20 paired-words was formed as the test stimuli. Words belonging to both the languages (Kannada and Hindi), having the same structure but different meanings (/mara/ meant 'tree' in Kannada and 'dead' in Hindi) were selected for the study.

2. *Selection of stimuli:* The most commonly used Kannada and Hindi words which formed homophones were chosen for the study. These words were given to three proficient native speakers and were judged by them as unfamiliar, familiar and highly familiar.

3. *Recording of stimuli:* The pairs of words were tape recorded by native

Kannada and Hindi speakers. They were asked to say the words as naturally as possible with an interval of 1 second between the words of a pair and an inter-stimulus interval of one minute.

Procedure

The final test material was presented to the participants in an individual set-up in a quiet environment. The stimuli were presented using a tape recorder. Before the actual administration of the test, one practice trial was provided to all the participants.

Task: The participants were asked to listen carefully to the pairs of words which were in two different languages that they knew and were asked to write the meaning of each word.

Responses: The responses of each individual were tabulated according to the number of correctly repeated words with correct meaning in each language (Kannada and Hindi).

Results

The data were tabulated and analyzed using SPSS version 16. Descriptive statistics (mean

Table 1. Mean and SD number of words correctly repeated with correct meaning

Age in Years	Group	Hindi words		Kannada words	
		Mean	SD	Mean	SD
10-12	Kannada	0.37	0.74	9.37	1.84
	Hindi	12.12	3.52	2.37	1.40
17-19	Kannada	9.25	4.71	8.37	5.55
	Hindi	8.16	3.65	8.33	3.98
35-45	Kannada	5.12	4.08	9.62	4.40
	Hindi	10.25	2.65	6.12	2.41
Total	Kannada	4.91	5.07	9.12	4.07
	Hindi	10.36	3.49	5.36	3.53

and standard deviation values) for both the groups of speakers for Kannada and Hindi words with respect to age group are shown in Table 1. The results are described in three stages.

Stage I: Comparison of age groups

One-way analysis of variance was carried out for the comparison of age groups separately for Kannada and Hindi speakers. Results revealed that for Kannada speakers, no significant difference was seen across ages for Kannada words at 0.05 level of significance. However, a statistically significant difference was observed { $F(2, 21) = 11.995, p < .001$ } among the three age groups for the retrieval of Hindi words. A post-hoc analysis (Duncan's test) revealed that all the age groups were significantly different from one another. Participants in the age group of 30 to 45 years performed significantly better than those in the age range of 10 to 12 years while those in the age range of 17 to 19 years performed significantly better than those in the age range of 30 to 45 years. This indicates that the number of words retrieved in L2 was more in teenagers followed by adults and then children.

In Hindi speakers no significant difference was seen for Hindi words across the age groups at 0.05 level of significance. However a statistically significant difference was seen across the age groups for Kannada words { $F(2, 19) = 9.140, p < .001$ }. From Duncan's post-hoc analysis it was evident that the performance of the participants in the age range of 10 to 12 years was significantly poorer than other two age groups.

Again the same trend was observed as in Kannada speakers, that is retrieval of words in L2 was more in teenagers followed by adults and then by children.

Stage II: Comparison between Kannada and Hindi speakers

To study the significant difference between Kannada and Hindi native speakers, independent t-test was carried out separately for all the three age groups within Kannada and Hindi words. It was observed that in group I which consisted of participants in the age range of 10 to 12 years, a significant difference was seen between the speakers for retrieval of both Hindi { $t(14) = 8.526, p < 0.000$ } and Kannada words { $t(14) = 9.230, p < 0.00$ }. Results showed that speakers performed significantly better in their native languages (refer Table 1). In group II (17-19 years), a significant difference between the speakers was not seen for both Hindi { $t(12) = 0.466, p > 0.05$ } and Kannada words { $t(12) = 0.016, p > 0.05$ }, i.e. their performance in L1 and L2 was statistically different. In group III, who had participants in the age range of 35-45 years, there was a significant difference for Hindi { $t(14) = 2.973, p < 0.010$ } but not for Kannada words { $t(14) = 1.970, p > 0.05$ }. Hindi speakers performed significantly better in Hindi words compared to Kannada speakers.

Stage III: Comparison between Kannada and Hindi words

To study the significant difference between

Kannada and Hindi words, the paired t-test was carried out separately for Kannada and Hindi native speakers in each age group. It was observed that in Kannada speakers there was a significant difference between retrieval of Hindi and Kannada words for participants in the age range of 10 to 12 years $\{t(7) = 11.906, p < 0.05\}$. However, statistically no significant difference in participants of other two age groups i.e. 17-19 years $\{t(7) = 0.246, p > 0.05\}$ and 35-45 years $\{t(5) = 0.056, p > 0.05\}$. Overall, the responses of the Kannada speakers revealed that their retrieval abilities were better for Kannada words than the Hindi words. This can be inferred from the mean value obtained (Table 1) for the retrieval of Kannada words which was greater than that obtained for Hindi words.

In Hindi speakers, a significant difference was observed for participants in the age range of 10 to 12 years $\{t(7) = 5.777, p < 0.001\}$ and those in the age range of 35 to 45 years $\{t(7) = 2.434, p < 0.05\}$, indicating that both the groups performed better in their native language. No significant difference

Kannada words.

The performances of the participants are also depicted for both Kannada (Fig.2.) and Hindi (Fig.3.) native speakers for all the age groups. It is evident from the graphs that both Hindi and Kannada native speakers in the age range of 10-12 years performed better in their native language. In case of Hindi speakers it was seen that they performed better in their L2 than compared to Kannada native speakers. However, statistically no significant difference was seen for the teenage group of Kannada and Hindi native speakers for both Kannada and Hindi words, i.e. they performed equally well in both the languages.

Discussion

The above results show that the retrieval of the meanings of the inter-lingual homophones is better in the native language (L1) in the age group of 10-12 years. This was seen both for Kannada and Hindi native speakers. These results are in

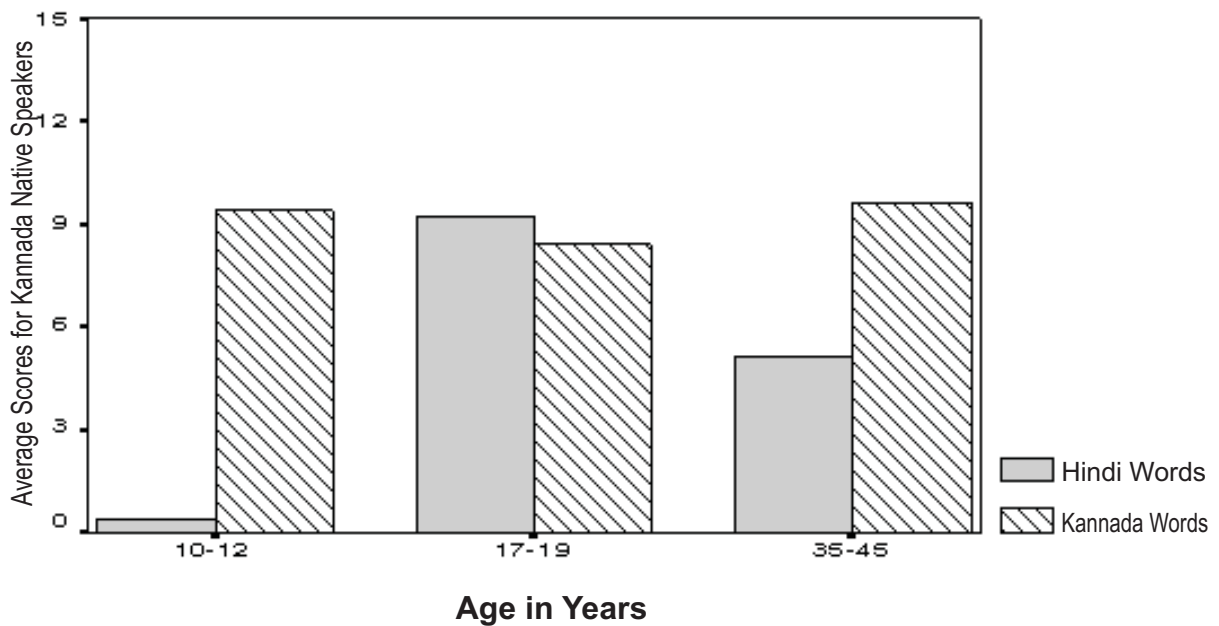


Fig.2. Inter-lingual homophone retrieval for Hindi and Kannada words in Kannada speakers

was seen between retrieval of Hindi and Kannada words for the age group of 17 to 19 years $\{t(5) = .056, p > .05\}$. It is evident from table-1 that the retrieval abilities were better for Hindi than

accordance with the studies done by Curtis (1978, 1999) stating that the bilinguals organize and retrieve words mainly in their dominant language. This difference of better retrieval

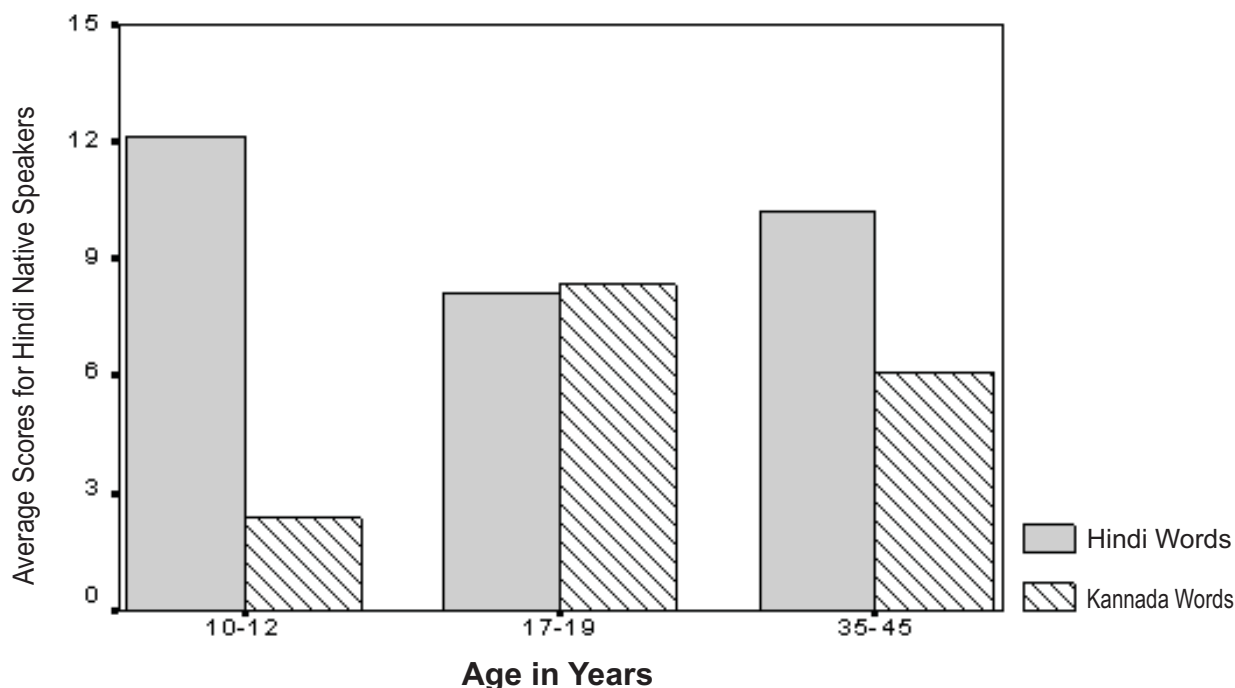


Fig. 3. Interlingual homophone retrieval for Hindi and Kannada words in Hindi speakers

abilities in the native language, i.e., L1 can also be well explained with the help of the Revised Hierarchical Model (Kroll & Stewart, 1994). Based on this model it can be assumed that the word-concept association has stronger links when the word belongs to L1 than in L2. This is indicative of the fact that L1 has a stronger base compared to L2 while processing inter-lingual homophones. Hence, the same model can be employed to explain the processing of homophones when presented through the auditory modality. The use of L1 in the younger age group more effectively is due to the lesser exposure to the second language. In the teenage period however a shared lexicon exists and further in adulthood the base of L1 becomes stronger compared to L2 and in order to establish a connection between the word and meaning in L2, the bilingual needs to access the conceptual level through L1.

The findings of the present study also suggest that children can process one language (L1) without the interference of the other (L2) effectively, giving the picture of two separate lexicons for both the languages. They show a selective lexical-access (i.e. only one language is stimulated at a time) and this is in accordance with

the findings of Gerard and Scarborough's (1989). However, the teenagers retrieved words equally well in both the languages. This shows that children were not having a sufficient command over the L2, while the teenagers showed an equal proficiency in both the languages i.e. L1 and L2. Similar findings have been advocated by Green (1986) who stated that when a person has reasonable command of both the languages, lexical items are subconsciously activated in both the languages. This group also exhibits a non-selective lexical access (i.e. both the languages are simultaneously accessed) as suggested by earlier researchers. The Kannada adult native speakers adults were more dominant in retrieving meaning in their native language (L1) but Hindi native speakers showed equal proficiency in retrieving words in both the languages. This difference in performance can be attributed to the greater exposure to Kannada (societal demands).

The results of the present study suggest that in an individual with a minimum formal education and exposure of 12 years or more appears to be equally proficient in L1 and L2. The results show that exposure plays an important role in becoming proficient in L2; however a formal education

becomes an added advantage. Thus, while assessing or designing language therapy the role of L1 and L2 may not be a variable especially at the semantic level.

Conclusion

Varied pattern of performance was seen across individuals of the three different age groups. Children were shown to have dominance in their native language and a similar trend was seen for adult speakers of Kannada. It was seen that although the exposure is same for all individuals, the teenagers (17-19 years) retrieve the meanings of words more efficiently for both the languages. The study has its implications in the assessment, diagnosis and planning of intervention programs. For teenagers, either L1 or L2 can be considered as a medium of instruction during remediation, but care should be taken about the choice of language used while planning out intervention for children and adults.

Acknowledgements: Authors would like to acknowledge:

Dr. Vijayalakshmi Basavaraj, Director, All India Institute of Speech and Hearing, for allowing me to carry out the study.

Ms. Swati for encouraging and helping in every step while carrying out this research work.

All the participants of the study for being so cooperative throughout the data collection.

Address for Correspondence: Ms. Ramya Maitreyee, All India Institute of Speech and Hearing, Manasagangothri, Mysore- 570006, India. E-mail: maitreyee_ramya@yahoo.co.in

References

- Beauvillian, C. & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language selective access. *Journal of Memory and Language*, 26, 658-672.
- Bybee, J.L. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- Costa, A. & Carmazza, A. (1999). Is lexical selection in bilinguals language-specific? Further evidence from Spanish-English & English-Spanish bilinguals. *Bilingualism, Language & Cognition*, 2, 231-244.
- Curtis, H. (1978). Recognition memory processes and language dominance in bilingualism. *Journal of Psycholinguistic Research*, 7, 25-34.
- Fromkin, V. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27-52.
- Gerard, L., & Scarborough, D. (1989). Language specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 305-313.
- Grainger, J. (1994). Le lexique bilingue: approches de la psychologie expérimentale. *Revue de Phonetique Appliquée*, 112/113, 221-238.
- Green, D. (1986). Control, activation and resource: a framework and a model for the control speech in bilinguals. *Brain and Language*, 27, 210-223.
- Hatch, E. (1983). *Psycholinguistic-a second language perspective*. Rowley, MA: Newbury House.
- Kimura, D. (1992). Sex differences in the Brain. *Scientific American*, September, 19-25.
- Kirsner, K. (1986). Lexical representation: Is a bilingual account necessary? In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives*. Hillsdale, NJ: Erlbaum.
- Kroll, J.F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Presentations. *Journal of Memory and Language*, 33, 149-174.
- Laufer, B. (1989). *A factor of difficulty in vocabulary learning: Deceptive transparency*. *AILA Review*, 6, 10-20.

- Lalor, E. & Kirsner, K. (2001). The representation of the "false cognates" in the bilingual lexicon. *Psychonomic Bulletin & Review*, 8, 552-559.
- Paradis, M. (1987). *The assessment of bilingual aphasia*. Hillsdale, NJ: Lawrence Erlbaum.
- Soudek, L.I. (1982). The mental lexicon in second language learning. Paper presented at the Thirteenth International Congress of Linguistics, Tokyo.
- Weinrich, U. (1953). *Languages in contact*. New York: Linguistics Circle of New York.
- Woutersen, M. (1996). Proficiency and bilingual lexicon. Paper presented at the Eighteenth Annual Conference of the American Association for Applied Linguistics. Chicago.
- Woutersen, M. (1997). *Bilingual word perception*. Nijmegen: Katholieke University. Nijmegen