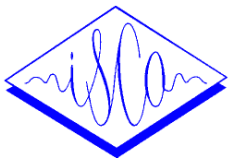

First International Conference on
Spoken Language Processing (ICSLP 90)
Kobe, Japan
November 18-22, 1990

LEXICAL MEMORY IN VISUAL AND AUDITORY MODALITIES: THE CASE FOR A COMMON MENTAL LEXICON

David B. Pisoni and Ellen E. Garber

Speech Research Laboratory, Indiana University
Bloomington, Indiana 47405 USA

ABSTRACT

This paper reports the results of a study designed to measure differences in familiarity for spoken and written words. Two sets of 450 English words were randomly selected from a computerized version of Webster's pocket dictionary. Four groups of subjects were presented with both lists of words for familiarity judgements. The first group (VV), saw all the words on each list presented visually; the second group (AA) heard all the words; the third and fourth groups (AV, VA) received one list visually and another list auditorily. Subjects rated the familiarity of each word using a seven-point scale. Correlations of the familiarity scores across both lists and modalities were very high. The mean ratings were not significantly different for visual and auditory groups. The absence of modality differences suggests that familiarity effects occur late in the processing system where information from the input modality converges on a common lexical store in long-term memory.

I. INTRODUCTION

For many years, the field of speech perception has been conspicuously isolated from the mainstream of research on spoken language processing that has addressed problems of word recognition, lexical access, sentence comprehension and discourse processing. This situation came about, in part, because the primary concern in speech perception research had been on the physical correlates of speech, that is, acoustic-phonetics [1]. Very little interest or attention was devoted to the interface between speech and more abstract levels of language processing. Fortunately, this situation has changed recently as more researchers have turned their atttention to somewhat broader issues that encompass not only the search for the acoustic cues to phoneme perception but the mapping problem between speech and the lexical representations of words in the listener's long-term memory. Many investigators realized that our understanding of how human listeners perceive words and understand fluent speech was very impoverished compared to the extensive literature on the perception of phonemes in isolated nonsense syllables. As researchers began to pursue these issues, new research strategies and techniques began to evolve to study speech perception and spoken word recognition [2].

For the last five years, we have been working with several computational techniques to learn more about the structural organization of spoken words in the mental lexicon and the role this organization may play in perception and memory [3]. To study these issues, we acquired several large lexical databases and carried out a number of analyses of the structural patterns of words and their organization. We computed similiarity spaces using several quantitative metrics that were designed to index the phonetic distance among sound patterns in English. And, we studied how these similiarity spaces affect perception in numerous behavioral experiments. One of these databases, Webster's Pocket Dictionary, contains 20,000 words. Each entry contains the standard English orthography for a word, a phonetic transcription, and special codes indicating the syntactic functions of the word. Frequency information was added to each entry when available by merging information from the Kucera and Francis word count. We also developed several ways of computing similiarity neighborhoods for words in this database so that quantitative measures of "lexical density" could be included with each entry. Numerous experiments over the last five years have demonstrated that the number and nature of words in a similiarity neighborhood strongly affects word recognition performance and memory in a variety of experimental tasks. In particular, neighborhood structure was shown to be an important determinant of the speed and accuracy of word recognition [4]. The number of words in a similiarity neighborhood and their relative frequencies have been shown to affect spoken word recognition performance in perceptual identification, lexical decision and naming tasks [4]. These same structural factors also affect encoding and retrieval of words from long-term memory [5].

In the course of carrying out this program of research, we also collected familiarity ratings from human observers for each of the 20,000 words in the Webster's database in order to have both objective frequency counts and subjective familiarity judgements for each entry [6]. Although many researchers have routinely assumed that objective frequency counts provide valid estimates of experienced frequency, other theorists have argued that this practice may be inappropriate because frequency counts substantially underestimate exposure to many words that are actually present in a listener's mental lexicon but were simply not included in the

corpora because of sampling problems. A number of years ago, Gernsbacher [7] reported that familiarity ratings from subjects were more accurate predictors of performance on several psycholinguistic tasks than published estimates of word frequency derived from objective word counts.

The familiarity data that we collected several years ago and routinely use in our experimental work were all obtained with visually presented materials. The question naturally arises as to whether these familiarity ratings can be generalized across modalities and whether they are appropriate for spoken as well as written words. Do these familiarity ratings reflect information that is, in some sense, modality-specific or do they index information about words that is modality-independent, information that presumably resides in some common mental lexicon that is shared across modalities? If familiarity judgements for the same set of words using visual and auditory presentation are comparable, or nearly so, one could argue that these results reflect access to or retrieval of information contained in a modality-independent lexicon rather than separate modality-specific lexicons that would be functionally autonomous for auditory and visual inputs. Results of this kind would not only be relevant to issues about the existence of a single modality-independent lexicon in language processing but they would also be extremely useful in resolving several long-standing problems concerning the locus of frequency effects in word recognition studies [8].

A number of theorists have argued that frequency effects occur very early in perceptual processing prior to lexical access and retrieval of the meaning of a word from long-term memory [9]. If familiarity effects are modality-independent and reflect access to information that is common to both auditory and visual inputs, the locus of these effects may therefore occur relatively late in perceptual processing presumably sometime after the two input modalities have converged on a common representation in lexical memory. To study this problem, we presented lists of words both visually and auditorily to separate groups of subjects to determine if the subjective familiarity ratings assigned to words would be dependent on the modality of input.

II. METHOD

2.1 Subjects

Ninety-six undergraduate students from introductory psychology courses served as subjects. Each subject attended two, fifty-minute sessions and received partial course credit for their participation. All subjects were native speakers of English who reported no history of a speech or hearing disorder at the time of testing.

2.2 Materials

Two lists of 450 words were constructed using a computerized lexical database derived from Webster's Pocket Dictionary. This database included frequency counts for 11,750 words that were obtained from the Kucera and Francis [10] word count as well as familiarity ratings from the earlier study by Nusbaum, Pisoni and Davis [6]. The two sets of words were randomly selected from the database with the restriction that each list contain 150 words with high familiarity ratings (5.1-7.0), 150 words with medium familiarity ratings (3.1-5.0), and 150 words with low familiarity ratings (1.0-3.0). For the auditory presentation, the words on each list were spoken in isolation by a male talker. For the visual presentation, the words were presented on a CRT using a video character generator that was controlled by a computer.

2.3 Procedure

Subjects were tested in groups of six or fewer in a sound treated room. Each subject was seated in an individual testing booth which was equipped with a GBC CRT monitor and a pair of TDH-39 headphones. Subjects were instructed to rate the subjective familiarity of each word using a seven point scale. The scale ranged from 7 (very familiar word) to 1 (word was unknown). Subjects recorded their familiarity judgements by pressing the appropriate button on a response box that was interfaced to a computer.

Each subject received two lists of words, one list per session. The lists were presented either visually (V) or auditorily (A) in each session. Twenty-four subjects participated in each of the four conditions, VV, AA, AV and VA. Within each condition, the order of presentation of the two lists was counterbalanced so that 48 subjects received List 1 first and another 48 subjects received List 2 first. Presentation of the stimuli was randomized for each group of subjects.

Presentation of the stimuli and collection of the data were controlled on-line by a PDP 11/34 minicomputer. In the auditory condition, subjects listened to the words over headphones. The onset of a light on the response box indicated the beginning of a new trial. In the visual condition, the words were presented in uppercase letters on a CRT monitor. A line of asterisks on the screen indicated the onset of a new trial. The word remained on the screen until the subject responded. In both the visual and auditory conditions, response times were measured from the onset of each stimulus until the subject initiated his/her response. Subjects had eight seconds in which to respond. At the end of the time period, the computer proceeded to the next trial.

III. RESULTS

Mean familiarity ratings were computed for each word in each modality of presentation. A summary of these results is shown in Figure 1. All four conditions displayed the same pattern of ratings across the three levels of familiarity. Analysis of variance confirmed the trends shown in the figure. Modality was not significant. Thus, the familiarity ratings were unaffected by mode of presentation. Not surprisingly, familiarity produced a highly significant effect ($p<.001$). No significant interactions were obtained.

The response latency data which is shown in Figure 2 displayed a similar trend across the three levels of word familiarity. Analysis of variance revealed significant main effects for modality and familiarity ($p<.01$). Overall, highly familiar words were responded to more rapidly than unfamiliar words and auditory presentation was slower than visual presentation. None of the interactions were significant.

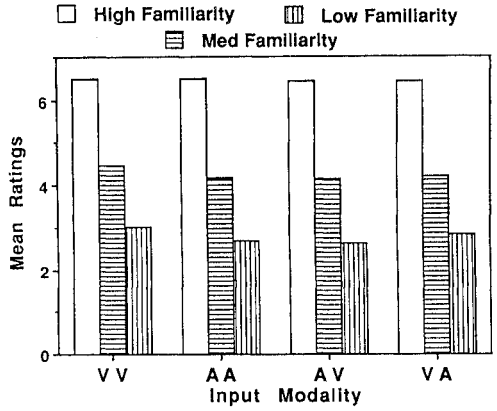

Fig. 1 - Mean familiarity ratings for high, medium and low words as a function of input modality.

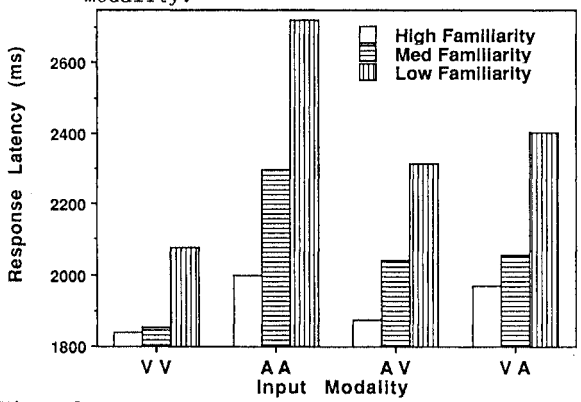

Fig. 2 - Response latencies for high, medium and low words as a function of input modality.

Figure 3 shows a scatter plot of the familiarity ratings for auditory and visual presentations of the two word lists. The correlations between the auditory and visual ratings were high and statistically reliable (r=+.93 and r=+.92, for Lists 1 and 2, respectively).

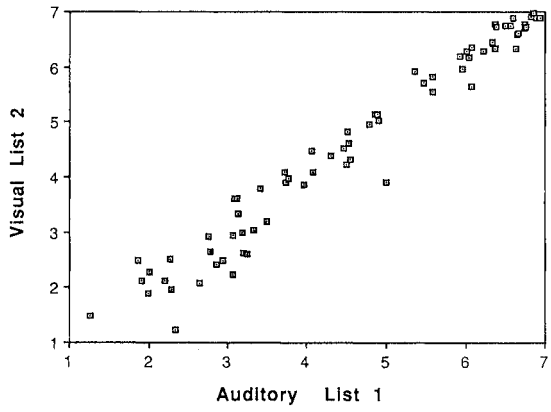

Fig. 3 - Scatterplot of familiarity ratings for auditory and visual presentation.

IV. DISCUSSION

The results of the present experiment clearly demonstrate that familiarity judgements are independent of input modality. The sensory modality used to present the word lists apparently did not affect subject's familiarity judgements in this task. Moreover, the correlations between the two word lists both within and between modalities were all positive and extremely high suggesting a common basis for the familiarity judgements that is independent of modality of input.

The results of this study have a number of implications for current research on word recognition and models of lexical access. First, in terms of our original goals, the present findings demonstrate that the familiarity ratings obtained in our earlier investigation, which used only visual presentation, can be generalized to spoken words with some degree of confidence. This is useful information to have when constructing materials for perceptual experiments and making generalizations across modalities.

Second, the absence of modality effects in this task raises several questions about the claims made by Forster [11] concerning the locus of frequency effects in word recognition experiments. Forster states that the frequency of a printed word may differ widely from the frequency of the spoken form of the same word. This assertion, among others, was used to motivate the proposal of three separate peripheral access files or bins in his search model. The entries in each bin are listed according to the frequency of occurrence of the word in the language. According to Forster's search model, a word is recognized by a two-step process. First, a description of the stimulus features of a word -- what Forster calls the "access code" is located by searching for an entry in one of the peripheral access files. When a match is found, the search process terminates and a pointer specifies the location of that word in the master file. The master file contains all the information that a listener has about a word whereas the peripheral access files contain only access codes and pointers to words in the master file. Thus, the peripheral access files are organized by both modality and frequency to permit efficent search. Within each modality-specific bin, the entries are arranged by frequency so that high-frequency words can be located earlier than low-frequency words. If we assume that word frequency somehow underlies or controls familiarity judgements through exposure to words in the language, then Forster's claim about differential frequency effects would appear to be incorrect because both frequency and familiarity effects are not influenced by input modality. A similar set of criticisms concerning frequency effects can be raised about Morton's [12] proposal of two input logogens to deal with modality-specific facilitation effects. In Morton's model, the logogens are not only frequency-sensitive but they are also modality-specific as well. Again, the present results question the need for proposing modality specific logogens to account for frequency effects.

Third, another closely related issue deals with the question of the precise locus of frequency effects in current models of word recognition. The well-known models of proposed by Morton [12], Forster [11] and Marslen-Wilson [9] all place the locus of frequency effects relatively early in perceptual processing prior to lexical access. The absence of any modality differences and the high correlations obtained in the present study within and across modalities suggests that frequency and familiarity effects occur fairly late in processing after the inputs have converged on a common representation rather than early on during search or activation as suggested by these investigators.

Finally, the present results provide additional evidence for a dissociation between between frequency and modality effects. Studies by Lee et al. [13] and Kirsner et al. [14] have found that frequency effects are largely independent of input modality. Their findings, taken together with the present results, suggest either a common modality-independent representation for words or the use of differential retrieval processes for modality-specific and frequency-specific information in long-term memory. Whatever the final explanation turns out to be, it is clear from the results of the present study that judgements of subjective familiarity for words are made on the basis of information that is not coded or indexed by input modality. Thus, there appear to be both common and modality-specific components to the mental lexicon. The extent to which these two sets of attributes can be accessed in various psycholinguistic tasks will depend largely on the specific requirements of the information processing task and the strategies adopted by the subject.

ACKNOWLEDGEMENTS

This research was supported by NIDCD Training Grant DC00012 and NIDCD Research Grant DC00111 to Indiana University in Bloomington. We thank Paul Luce for his help and advice.

REFERENCES

[1] A.M. Liberman, et al., "Perception of the Speech Code," Psychol.Rev., Vol. 74, pp.431-461, 1967.

[2] D.B. Pisoni, "Speech Perception: Some New Directions in Research and Theory," J.Acoust.Soc.Am., Vol. 78, pp.381-388, July, 1985.

[3] D.B. Pisoni, "Speech Perception, Word Recognition and the Structure of the Lexicon," Spch.Comm., Vol. 4, pp.75-95, August, 1985.

[4] P.A. Luce, "Neighborhoods of Words in the Mental Lexicon," Res. on Spch.Percep. Tech.Report No. 6, Indiana Univ., pp.1-91, December, 1986.

[5] S.D. Goldinger, et al., "On the Locus of Talker Variability Effects in Recall of Spoken Word Lists," J.Exp.Psychol.: Learn., Mem. & Cog. (In Press).

[6] H.C. Nusbaum, et al., "Sizing up the Hoosier Mental Lexicon: Measuring the Familiarity of of 20,000 Words," Res. on Spch.Percep. Prog.Report No. 10, Indiana Univ., pp.357-376, December, 1986.

[7] M.A. Gernsbacher, "Resolving 20 Years of Inconsistent Interactions between Lexical Familiarity and Orthography, Concreteness, and Polysemy," J.Exp.Psychol.: Gen. Vol. 113, pp.256-281, 1984.

[8] D.A. Balota & J.I.Chumbley, "Are Lexical Decisions a Good Measure of Lexical Access? The Role of Word Frequency in the Neglected Decision Stage," J.Exp.Psychol.: Hum.Percep. & Perform. Vol. 10, pp.340-357.

[9] W.D. Marslen-Wilson, "Function and Process in Spoken Word Recognition: A Tutorial Review," in H. Bouma & D.G. Bouwhuis (Eds.) Attention and Performance X.: Control of Language Processes, Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 125-150, 1984.

[10] F. Kucera & W. Francis, "Computational Analysis of Present Day American English," Providence, R.I., Brown University Press, 1967.

[11] K.I. Forster, "Accessing the Mental Lexicon," in R.J. Wales & E.C.T. Walker (Eds.), New Approaches to Language Mechanisms, Amsterdam, North-Holland, pp.257-287, 1976.

[12] J. Morton, "Facilitation in Word Recognition: Experiments Causing a Change in the Logogen Model," in P.A. Kolers, M.E. Wrolstad & H. Bouma, Processing Visible Language, New York, Plenum Press, pp.259-268, 1979.

[13] A.T. Lee, et al., "Sensory Modality and the Word-Frequency Effect," Mem. & Cog. Vol. 6, pp.306-311, 1978.

[14] K. Kirsner, et al., "Common and Modality-Specific Processes in the Mental Lexicon," Mem. & Cog. Vol 11, pp.621-630, 1983.