

VOCAL COMMUNICATION OF EMOTION:

A PROGRAM OF RESEARCH

by Ken W. McCluskey

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE
STUDIES IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

UNIVERSITY OF MANITOBA
WINNIPEG, MANITOBA, CANADA

OCTOBER, 1979

VOCAL COMMUNICATION OF EMOTION:

A PROGRAM OF RESEARCH

BY

KENNETH McCLUSKEY

A thesis submitted to the Faculty of Graduate Studies of
the University of Manitoba in partial fulfillment of the requirements
of the degree of

DOCTOR OF PHILOSOPHY

© 1980

Permission has been granted to the LIBRARY OF THE UNIVER-
SITY OF MANITOBA to lend or sell copies of this thesis, to
the NATIONAL LIBRARY OF CANADA to microfilm this
thesis and to lend or sell copies of the film, and UNIVERSITY
MICROFILMS to publish an abstract of this thesis.

The author reserves other publication rights, and neither the
thesis nor extensive extracts from it may be printed or other-
wise reproduced without the author's written permission.

ABSTRACT

The present program of research consisted of three studies. The purpose of Study I was to examine and compare the ability of Canadian and Mexican subjects to identify the emotional content of vocal expressions. Canadian and Mexican female teachers, speaking any words they wished in their respective languages (i.e., English and Spanish), simulated the emotions of happiness, sadness, love, and anger. After the initial recording, speech samples were arranged in random order and then rendered unintelligible by means of an electronic filtering device (which removed semantic content while leaving intact the tonal qualities of speech). The filtered vocal expressions were played to Canadian and Mexican subjects 5, 9, 13, 17, 25, 45, and 65 years of age. From 5 through 25, there was a progressive increase with age in ability to identify the emotion expressed. Ability did level off, however, and eventually even decreased (such that the 65-year-olds performed less accurately than the other adult groups). Overall, the Mexican subjects were significantly more sensitive to tone than their Canadian counterparts. However, subjects from both countries judged the speech samples from Mexican teachers more accurately than those from the Canadian speakers. As well, there was a response bias, with younger subjects in both cultures emitting more negative

than positive responses when judging the vocal expressions. Adult subjects, on the other hand, showed no tendency to respond with a preponderance of negative judgments.

In Study II, recordings were made of bilingual Canadian and Mexican female teachers reciting congruent and contradictory messages in both English and in Spanish. In the congruent condition, the verbal and tonal elements of the communication were consistent. With the contradictory speech, however, the tone of voice was incongruent with the words spoken. When the appropriate speech samples (i.e., English for the Canadian subjects and Spanish for the Mexicans) were played to Canadian and Mexican boys aged 5, 9, 13, and 17, it was found that there was a clear speech effect. That is, children at all age levels in both cultures reacted more negatively to contradictory than to congruent messages. Also, there was an age effect, with the younger children in both cultures interpreting the contradictory messages more negatively than the older ones. Further, a culture effect was noted, such that the Mexican children viewed the contradictory speech significantly more negatively than did their Canadian counterparts.

In Study III, variables other than culture were the major concern. As a consequence, only Canadian subjects participated. The congruent and contradictory speech samples recorded in English by the Canadian teachers for the second study were again used here. In this case, however, all the

congruent and contradictory vocal expressions were arranged in random order on a single tape and played to "normal" and "disturbed" boys of different ages (i.e., at the grade 2, the grade 4, and the grade 6 levels). Significant effects were found for type of speech and age: subjects reacted more negatively to contradictory than to congruent speech, and younger children responded more negatively to contradiction than did older ones. As well, contradictory, double-bind type speech seemed to have more of an impact on the "disturbed" boys than on the "normals." The fact that "disturbed" children reacted more negatively to contradictory communications was interpreted as being generally supportive of the double-bind hypothesis (after it had been redefined and softened in certain respects).

ACKNOWLEDGEMENTS

I'd like to extend sincere thanks to my academic advisor, Dr. John McIntyre, for all his assistance in the planning and design of this dissertation. Essentially, Dr. McIntyre picked me up as a student when no one else wanted me, and provided the help and support I needed to finish up the Ph.D. He continually made meaningful and constructive suggestions to help me along, giving very freely of his own time all the while. Indeed, without his supervision (which, paradoxically enough, managed to be flexible and structured simultaneously) this thesis would probably have remained unwritten.

Considerable thanks are also extended to Dr. Ross Hartsough and Dr. F.L. Marcuse not only for serving on the committee, but also for offering excellent courses in the nonverbal area (which served to shape and influence my thinking long before I undertook the research at hand). Their many suggestions regarding style and form were much appreciated. As well, I must thank Dr. Dan Albas for all the time spent "digging up" articles on my behalf ... his input with respect to the cross-cultural literature was especially helpful. Thanks also go to Dr. E.H. Schludermann for his suggestions during the preliminary oral.

Dr. Walter J. Lonner took time from his busy schedule to serve as external examiner for this dissertation. I very much appreciate his constructive comments and suggestions, and I feel the thesis benefited a great deal from his involvement. His many books in the cross-cultural realm were also extremely helpful.

At a more personal level, I want to thank my family for all their encouragement and moral support during this "period of suffering." Everyone at home knows Mom was involved in my university career every bit as much as I was. Her affectionate "nagging" kept me going ... I'm likely the only individual who completed a doctorate for the sole purpose of keeping his mother happy. Dad, of course, was always willing to help, and I've lost count of all the little "errands" he took care of so that I could work on without interruption. Now that I've finished up, he can finally stop playing hockey with the pros. I musn't forget Jill and Mark--who both like to see their names in print--for their involvement and patience. Due to clever planning (years of procrastination) on my part, Jill and I will be graduating together. In any case, likely no one at U of M has owed more to his family -- it was nice to have them in "on the kill." It was also nice to have Kathryn--a very special girl in my life--along as well. After the oral she noted (in her usual self-congratulatory fashion) that: "You never could have done it without me!" That facetious comment just might be true.

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	i
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	x
LIST OF FIGURES	xi
STUDY I	1
INTRODUCTION	2
METHOD	9
SUBJECTS	9
APPARATUS	10
STICK FIGURES	10
SPEECH SAMPLES	12
PROCEDURE	14
RESULTS	15
DISCUSSION	22
STUDY II	28
INTRODUCTION	29
METHOD	35
SUBJECTS	35
CONTRADICTORY AND CONGRUENT SPEECH SAMPLES	35
PROCEDURE	37
RESULTS	38
DISCUSSION	43

	<u>Page</u>
STUDY III	46
INTRODUCTION	47
METHOD	51
SUBJECTS	51
SPEECH SAMPLES	52
PROCEDURE	53
RESULTS	54
DISCUSSION	57
SUMMARY	60
APPENDIX I: REVIEW OF THE LITERATURE	63
NONVERBAL COMMUNICATION	64
CHANNELS OF COMMUNICATION	67
The Verbal Channel	67
The Vocal Channel	69
The Visual Channel	72
Nonverbal Channels in Combination	76
EMOTION	78
Coding the Vocal Channel for Emotional Content	80
JUDGMENT OF VOCAL ASPECTS OF SPEECH BY ADULTS	85
Recitation of Nonmeaningful Material	85
Recitation of Standard Content	87
Filtered Speech	88
JUDGMENT OF VOCAL ASPECTS OF SPEECH BY CHILDREN	91

	<u>Page</u>
CROSS-CULTURAL DIFFERENCES IN THE PERCEPTION OF THE EMOTIONAL CONTENT OF NONVERBAL MESSAGES	98
Cultural Universals?	100
The Effects of Culture	102
Cultural Effects in Childhood	110
THE DOUBLE-BIND HYPOTHESIS	112
The First Decade of Double-Bind Research	116
The Second Decade	125
Research With Children	128
A Redefinition	133
APPENDIX II: REFERENCES	135
APPENDIX III	165
APPENDIX IV	167
APPENDIX V	170
APPENDIX VI	178
APPENDIX VII	180
APPENDIX VIII	182
APPENDIX IX	184
APPENDIX X	186
APPENDIX XI	188
APPENDIX XII	190
APPENDIX XIII	192
APPENDIX XIV	194
APPENDIX XV	196
APPENDIX XVI	198
APPENDIX XVII	200

	<u>Page</u>
APPENDIX XVIII	202
APPENDIX XIX	204
APPENDIX XX	206
APPENDIX XXI	208
APPENDIX XXII	210
APPENDIX XXIII	212
APPENDIX XXIV	214
APPENDIX XXV	216
VITA	218

LIST OF TABLES

	<u>Page</u>
1. Mean number of negative responses by Canadian and Mexican subjects at successive age levels	20
2. Individual, total, and mean correct responses (scored against tone) for Canadian and Mexican subjects at the various age levels in the contradictory condition	42

LIST OF FIGURES

	<u>Page</u>
1a. Stick figures representing the four emotional states	11
1b. Mean number of correct identifications of Canadian and Mexican speech samples by Canadian and Mexican subjects at successive age levels	16
2a. Mean number of negative responses emitted to contradictory and congruent speech as a function of age and culture	39
3a. Faces (and numerical weightings) representing the emotional reactions to the speech samples	53
3b. Mean weighted scores to congruent and contradictory speech as a function of age and type of subject	55

STUDY I

PERCEPTION OF THE EMOTIONAL CONTENT OF SPEECH BY CANADIAN
AND MEXICAN CHILDREN, ADOLESCENTS, AND ADULTS

Any spoken communication between two or more people involves both verbal and nonverbal elements. Indeed, much of the research in the area of communication is based upon the notion that speech is composed of two distinct channels: the verbal and the vocal. Supposedly the *verbal* channel is composed solely of sound patterns resulting in syllables, words, phrases, sentences, and other linguistic forms conveying the semantic meaning of speech. If the words spoken in a particular message are written down, the verbal aspects of the communication have, in effect, been isolated. In fact, the verbal information contained in such written messages is what is most commonly examined by traditional content analytic techniques (cf. Barcus, 1959; Carney, 1972; Holsti, 1969).

The features of the voice itself, irrespective of the semantic meaning of the words spoken, are thought to come in on what is called the *vocal* channel. In other words, the vocal channel can be thought of as simply the tonal qualities of speech. Soskin and Kauffman (1961) were among the first to suggest that this channel carried potential affective information. They felt that the vocal channel was the "carrier" upon which verbal content is superimposed, and that the cues to emotion supposedly reside within this "carrier."

A number of investigators have approached the study of communication of emotion in human speech by looking at the vocal channel alone. In their classic investigation, Davitz and Davitz (1959a) tested the hypothesis that emotion can be reliably communicated by the nonverbal aspects of speech. In

an effort to remove the semantic meaning from communication, they had speakers recite parts of the alphabet while attempting to simulate various emotions. Adult subjects then made judgments as to what emotion the speaker was intending to communicate.

Electronic filters have also been used to remove semantic meaning before playing the speech samples to judges (Rogers, Scherer, & Rosenthal, 1971; Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979; Soskin & Kauffman, 1961; Starkweather, 1956). This latter technique may be preferable in that subjects are actually allowed to speak words, rather than trying to express emotion while saying the alphabet. Furthermore, the filtering method can be employed in studies concerned with real-life situations. For example, taped speech samples obtained from therapy sessions, telephone conversations, radio, television, and so on can be filtered and played to subjects; acting need not be involved. In any case, regardless of the method used, the literature on vocal expression of emotion, reviewed by Starkweather (1961), Davitz and Davitz (1961), Kramer (1963), Davitz (1964), Mahl and Schulze (1964), Duncan (1969), Mehrabian (1972) and Rosenthal, Hall, DiMatteo, Rogers, and Archer (1979) indicates that adults in general are able to identify the affective content of both nonfiltered and filtered vocal expressions.

Although it has been shown that adults can recognize the emotional content of speech, there is a decided paucity of research concerning the development of sensitivity to the

emotional meaning communicated via tone of voice (Nash, 1970). In one of the few developmental studies in the area, Gates (1927) assessed the ability of children from grades three to eight to judge the emotional content of portions of the alphabet recited to express each of a number of different emotions. She found that there was a progressive increase with age in ability to identify correctly the emotional content of the voice samples.

The next study with children along these lines was performed almost four decades later by Dimitrovsky (1964), who attempted to isolate the vocal channel of speech by holding the verbal channel constant. In her speech samples, speakers were asked to recite the same standard paragraph while simulating the emotions of happiness, sadness, love, and anger. When these samples were played to children aged five through twelve, it was likewise found that the ability to judge vocal expression of emotion was positively related to age. Fenster (1967), using a similar standard content (nonfiltered) method, and McCluskey (1974), employing both standard content and filtered speech samples, also found that the ability to identify emotion in vocal expressions improves with age.

Additionally, in terms of frequency of response, it is interesting to note that while adults do not seem to exhibit a response bias when making emotional judgments of speech samples (Davitz, 1964), there is some indication that children do tend to emit significantly more negative ("sad" and "angry") than positive ("happy" and "loving") responses (Dimitrovsky,

1964; McCluskey, 1974). Since her data indicated that youngsters respond with a greater number of negative responses, Dimitrovsky (1964) concluded that children are characteristically more negative in their judgment of vocal expressions than are adults. Dimitrovsky's findings and conclusions in this regard, however, have not necessarily been accepted by all other investigators. Fenster (1967) and Fenster and Goldstein (1971), for example, contended that both children and adults show no negative response bias when listening to a standard, content-free sentence.

To date, most developmental studies have dealt with American or Canadian subjects. It would seem useful, then, to extend the scope of investigation and approach the development of emotional sensitivity from a cross-cultural perspective. In the research at hand, Canada and Mexico were chosen as the comparison cultures, largely because they are so radically different from one another. Despite the fact that they are both technically North Americans, pronounced differences in "emotional style" seem to exist between Canadians and Mexicans. From an anecdotal point of view, Canadians often describe Mexicans as being "Latin blooded" and "highly emotional." Conversely, Canadians and other Anglo races are typically stereotyped by Mexicans as being cold, aloof, and "mechanical" (cf. Lewis, 1961).

At an empirical level, there is evidence to indicate that emotional expression does in fact vary from culture to culture. Specifically, with respect to nonverbal communicative

patterns, clear cultural differences exist in terms of size of personal space zones, posture and angle of orientation in interpersonal interaction, volume and tone of voice, and amount of gesturing, eye-contact, and bodily-contact (cf. Albas, McCluskey, & Albas, 1976; Argyle & Dean, 1965; Birdwhistell, 1963; Hall, 1959, 1964, 1966, LaBarre, 1947; Lacroix & Rioux, 1978; McCluskey & Albas, 1978; Mehrabian, 1972; Sommer, 1969; Watson & Graves, 1966).

Of particular interest is a study by Baxter (1970) which showed that Mexican children, adolescents, and adults had smaller personal space zones (and engaged in more bodily-contact) than either black or white Americans. The nonverbal styles of expressing emotion were clearly different between the Mexican and American subjects, and differences found in childhood tended to persist into adult life. Further, in a pilot study examining sensitivity to tone of voice, Mexican children aged six through 11 were found to identify the emotional content of filtered speech more accurately than Canadian children of corresponding age (McCluskey, Albas, Niemi, Cuevas, & Ferrer, 1975).

For the purposes of the present study, the issue is whether or not the subtle nonverbal differences between the Canadian (Anglo) and Mexican (Latin) cultures extend to even the tonal level. Although there is some empirical support for the intuitive notion that communication of affect differs in the two cultures (cf. Argyle, 1972; Mehrabian, 1972), the question as to whether there are cross-cultural differences

in sensitivity to the emotional content of speech has not been adequately explored. Indeed, there seems to be only one study addressing itself to this issue (McCluskey, Albas, Niemi, Cuevas, & Ferrer, 1975). The major aim of the present research, then, was to compare the ability of Canadian and Mexican subjects of various ages to judge the affective content of filtered vocal expressions recorded by Canadian (English-speaking) and Mexican (Spanish-speaking) communicators.

A complete literature review (Appendix I) and list of references (Appendix II) are included elsewhere in this thesis. As mentioned earlier (and in more depth in the detailed literature review), virtually all research involving recitation of nonmeaningful material and standard content has shown that children are able to judge emotion in speech more accurately as they grow older (Dimitrovsky, 1964; Fenster, 1967; Gates, 1927). Also, with filtered speech where there were no words or verbal content for the children to "hook onto," the same general trend has been reported (McCluskey, 1974). In light of such research, it was hypothesized that:

- (1) there would be a progressive increase with age in children's ability to identify the emotional content of filtered speech in both the Canadian and Mexican cultures.

As well, since no complete developmental comparison between Canadians and Mexicans has been undertaken, it was decided to present speech samples to Canadian and Mexican subjects of various ages (i.e., to children, adolescents, young adults, the middle aged, and the elderly). Keeping

in mind Baxter's (1970) notion that cultural responses learned early in the childhood years tend to persist into adulthood, it was felt that any cross-cultural differences in sensitivity to tone might well perseverate through the entire life cycle. It is known that certain differences exist in the communication of emotion in the Anglo and Latin cultures (Argyle, 1972; Baxter, 1970; Hall, 1959; McCluskey, Albas, Niemi, Cuevas, & Ferrer, 1975; McCluskey, Niemi, & Ferrer, 1973; Mehrabian, 1972). Given what research there is in the area, and assuming that the widespread popular belief that Mexicans are "extremely" emotional may perhaps have some basis in fact, the most logical and parsimonious expectation was that:

(2) Mexican subjects at all age levels would identify the emotional content of the filtered speech more accurately than their Canadian counterparts.

Finally, remembering Dimitrovsky's (1964) contention that children tend to be characteristically more negative in their responses to speech samples than adults, it was hypothesized that:

(3) while adults would not show a response bias in either direction, children in both cultures would make more negative ("sad" and "angry") judgments than positive ("happy" and "loving") ones.

METHOD

SUBJECTS

Sixty boys from the Lord Selkirk School Division #11, Selkirk, Manitoba, Canada, and 60 from schools in Colonia del Valle, Mexico City, Mexico, served as subjects. More specifically, 15 boys from each country at each of the ages of 5, 9, 13, and 17 took part. No subject at any age level was within two months of a birthday at the time the study was run.

As well, 15 male teachers at each of the *mean* age levels 25, 45, and 65 were selected from both Manitoba and Mexico City schools. Due to the difficulty in finding enough individuals at each age level, subjects ranged (plus or minus two years) around the mean ages listed above. That is, the ages actually ranged from 23-27, 43-47, and 63-67. All subjects in the 65-year-old group had retired from active teaching (but were contacted through their former schools).

Every subject at each age level in both countries had an I.Q. score that fell in the normal range (90-110) on the PPVT (Peabody Picture Vocabulary Test), QT (Quick Test), Slosson Intelligence Test, or on any group intelligence test. Also, the hearing of every subject participating in the study was tested within normal limits on a Bell Tone Audiometer.

Since social class seems to be a factor influencing ability to encode and decode nonverbal information (Hore, 1970; Knight & Kagan, 1977; Quay, Mathews, & Schwarzmuller, 1977;