12. CLINICAL ASSESSMENT AND SIMULATION

Chaired by Doris Aaronson, *New York University*

Computerized speech recognition in psychological assessment: A Macintosh prototype for screening depressive symptoms

GERARDO M. GONZÁLEZ

California State University, San Marcos, California

Computerized speech-recognition technology holds strong promise for psychological assessment. This paper focuses on a computerized speech recognition application for screening depressive symptomatology. A Macintosh-based prototype has been developed that uses the Center for Epidemiological Studies-Depression scale (CES-D). The prototype is a HyperCard stack interfaced with the Voice Navigator II speech recognition application. The "talking" program represents a viable depression screening tool that is fully voice operated by the respondent. A pilot study will assess the feasibility and acceptability of the computerized and written versions of the CES-D with a nonclinical adult population. A counterbalanced design will test for order effects and analyze the psychometric equivalence of the two methods. The limitations and future directions for speech-recognition applications in psychological assessment are discussed.

Clinical depression is one of the most debilitating affective disorders. The epidemiological literature reveals that it is the most common mental disorder in the United States. It affects 6%–7% of the general adult population, and between 10% and 25% of the population report significant depressive symptoms (Robins et al., 1985; Weissman, Bruce, Leaf, Florio, Holzer, 1991). However, only 20%–25% of the individuals who meet the criteria for clinical depression actually seek mental health professionals. Approximately 75% of clinically depressed individuals visit a medical care provider for relief of their depressive symptoms (Shapiro et al., 1984). Many medical patients are not screened for depression because most nonpsychiatric physicians are not adequately trained to evaluate it. As many as 30% of the primary care medical patients seen by physicians have been found to be clinically depressed. Yet more than half are inaccurately diagnosed for depression (Broadhead, Clapp-Channing, Finch, & Copeland, 1989; Pérez-Stable, Miranda, Muñoz, & Ying, 1990). Thus, clinical depression poses a serious mental health concern, because it remains largely undetected and untreated.

Computers are viable tools as automated interviewers for reliably assessing depression (Kobak, Reynolds, Rosenfeld, & Griest, 1990). Studies have revealed that depressed patients readily accept computer interactive interviewing (Carr, Ancill, Ghosh, & Margo, 1981; Carr, Ghosh, & Ancill, 1983). In fact, computerized interviewing may be more accurate than human interviewing in obtaining patient disclosure of suicide risk and predicting suicide attempts (Levine, Ancill, & Roberts, 1989). Selmi, Klein, Griest, Sorrell, and Erdman (1990) found that a computer-administered cognitive-behavioral program was as effective as a trained therapist in the treatment of mild to moderately depressed outpatients.

SPEECH RECOGNITION

Computerized speech recognition holds strong promise for psychological assessment. The potential is particularly significant for use with the disabled (Noyes, Haigh, & Starr, 1989; Tronconi, Billi, Boscarelli, Graziani, & Susini, 1989) and preliterate individuals who cannot functionally read or write. Many of these individuals are depressed or are vulnerable to depression as a result of greater psychosocial and emotional stressors. Conventional paper-and-pencil questionnaires are not a suitable means for measuring depression in such people. However, they are less likely to obtain services that largely require standard written assessment. Thus, the use of speech-

The author wishes to acknowledge the support of the California State University, San Marcos, Faculty Affirmative Action Grant, which provided released time for the development of the Macintosh prototype. Correspondence should be addressed to G. M. González, Psychology Program, California State University, San Marcos, CA 92096.

computerized screening provides greater accessibility to assessment and treatment services.

Richards, Fine, Wilson, and Rogers (1983) developed a voice recognition system for administering the Minnesota Multiphasic Personality Inventory (MMPI) to disabled patients with limited hand function. The system visually displayed the MMPI items and the patient verbally responded. The results indicated that there were no significant differences between the profiles generated by the computerized and written methods.

Muñoz, González, and Starkweather (1991) at the University of California, San Francisco, successfully pilot tested an IBM-compatible speech recognition "talking" prototype with English and Spanish-speaking depressed medical patients. The program verbally presented the 20 items of the Center for Epidemiological Studies—Depression scale (CES-D), recognized the patient's responses, and generated a report of the patient's level of depressive symptoms. The results suggested that the computerized CES-D was feasible to administer to a literate depressed medical patient sample. Both the computerized and written versions were found to be psychometrically equivalent. Moreover, a majority of the patients from both language groups preferred the computerized version over the written version.

THE PROTOTYPE

The Macintosh prototype represents a viable "talking" depression-screening application that is fully voice operated by the respondent. The measure is the Center for Epidemiological Studies—Depression (CES-D) scale (Radloff, 1977). The CES-D is a 20-item self-report instrument that assesses the frequency of a respondent's feelings, behavior, and mood during the week preceding the report. The respondent selects from among four choices (0 = less than 1 day, 1 = 1–2 days, 2 = 3–4 days, and 3 = 5–7 days). The scale includes four reverse-scored items phrased in a nondepressive direction. A total score, ranging from 0 to 60, is obtained by summing the weighted scores. A score of 16 or greater indicates a high level of depressive symptoms (Weissman, Sholomskas, Pottenger, Prusoff, & Locke, 1977).

The prototype utilizes any Macintosh computer with audio capabilities, such as a Macintosh Plus or later model. To develop the prototype HyperCard 2.1, a scripting tool (Claris, 1991) was interfaced with the speech recognition application Voice Navigator II (Articulate Systems, 1990). The program runs on System 6.0.7 or later and requires at least 2 MB of RAM.

The HyperCard-based prototype records and plays back stored digitized *speechfiles*. The computer program instructions and CES-D items were recorded as prompts for the user. Voice Navigator II uses speaker-dependent speech recognition. The user trains the computer to recognize the CES-D choices by creating a *voice template*. For each phrase, the user's voice characteristics are con-

Table 1
Summary of the Flow of the Macintosh
Speech Computerized Prototype

1. Interviewer enters user data
2. Computerized introduction to the user
3. User voice template training
4. User voice template testing
5. Computerized instructions for completing the CES-D
6. User-Computer interaction
 - (a) Computerized presentation of each CES-D item
 - (b) User states choice and the computer responds
7. Computerized closing remarks to the user
8. Interviewer scores and saves the CES-D results
9. Interviewer prints an interpretive report

verted into a digitized template and stored in memory. The program generates algorithms to subsequently recognize the spoken response by comparing it with the template. When a match is recognized, the program registers and scores the response. Thus, spoken commands eliminate the user's need for keyboard input.

The Macintosh prototype has been designed to perform several important operations. The interviewer first launches the program, enters the user's data, and prepares the user for interacting with the computer. The "talking" HyperCard stack uses a series of cards to play a brief introduction and provide instructions for training a template. The spoken instructions are also graphically displayed on the computer screen with text, buttons, and icons. The user is asked to repeat each CES-D choice several times to create a template. After template training, several cards test the recognition accuracy of the template. Following a successful test, the program presents the instructions for completing the CES-D. The interaction between the computer and user involves the presentation of a CES-D item card and the user's spoken response. The interaction continues until each item is answered. After all the items are completed, the program makes closing remarks. The interviewer scores and saves the results in a data file and stores a brief interpretive report. The interviewer has the option to print the report at a later time by using any word-processing program. A summary of the basic flow of the program is described in Table 1.

A pilot study will assess the feasibility and acceptability of the computerized and written versions of the CES-D with an English-speaking nonclinical adult population. A counterbalanced design will test for the effects of the order of administration for both methods. The psychometric properties of the two methods will be analyzed for determining equivalence by examining means, variances, inter-item consistency (coefficient alpha), and immediate test-retest reliability (Matarazzo, 1986; Wilson, Genco, & Yager, 1985). Computerized testing has been known to elicit anxiety and invalidate results (Canoune & Leyhe, 1985). Thus, the Computer Anxiety Rating Scale-Form C (Rosen, Sears, & Weil, 1987) will be utilized to explore the relationship of computer anxiety with the acceptability of the CES-D method.

FUTURE DIRECTIONS

Speech recognition presents numerous potential psychological assessment applications. Many psychological tests can be programmed for speech recognition in any spoken language, and any psychological questionnaire that has a discrete choice format may be converted to a speech recognition method. "Talking" applications have huge speechfile storage demands, which limits the size of the program. For portability on high-density diskettes, the size of the prototype was limited to 1.3 MB. However, the advent of optical and compact disks will overcome the storage limitations associated with numerous speechfiles.

Speech recognition can be used with portable computers and taken into the field. Newly developed software-only applications will eliminate the need for external input devices. Speech-computerized screening with a telephone interface is also possible, which is advantageous, because telephones are more commonly available than computer terminals and less anxiety provoking for some users. The portability of computerized screening will allow greater accessibility to clinical and nonclinical populations of people with mild to severe disorders in a variety of field settings.

As noted, speaker-dependent speech recognition requires the user to complete a template training session. For the prototype, the user generally takes 3-5 min to create a template of the CES-D choices. Thus, the computerized version typically takes longer than the written version. Furthermore, speech-recognition difficulties are generally related to the system's sensitivity to precise spoken repetition. A change in the user's tone, pitch, or inflection will result in decreased phrase recognition (Noyes, Haigh, & Starr, 1989). However, the recent breakthrough in speaker-independent continuous speech recognition will eliminate the need for voice template training.

Speech-recognition applications provide viable tools for enhancing and expanding mental health services to more people of any language. The objective is to increase and disseminate services that are otherwise limited by the lack of adequate funding, the lack of trained staff, or the lack of staff fluent in other languages. Computerized speech-recognition technology will extend the capabilities of the mental health professional by reaching more people who would otherwise continue to suffer debilitating depressive disorders. A reliable and valid speech recognition screening method will increase the proportion of accurately identified cases with high depressive symptoms and will lead to appropriate referral, prevention, and treatment of clinical depression.

REFERENCES

ARTICULATE SYSTEMS (1990). *Voice Navigator II* [Computer program]. Cambridge, MA: Articulate Systems, Inc.

- BROADHEAD, W. E., CLAPP-CHANNING, N. E., FINCH, J. N., & COPELAND, J. A. (1989). Effects of medical illness and somatic symptoms on treatment of depression in a family residency practice. *General Hospital Psychiatry*, **11**, 194-200.
- CARR, A. C., GHOSH, A., & ANCILL, R. J. (1983). Can a computer take a psychiatric history? *Psychological Medicine*, **13**, 151-158.
- CARR, A. C., ANCILL, R. J., GHOSH, A., & MARGO, A. (1981). Direct assessment of depression by microcomputer: A feasibility study. *Acta Psychiatrica Scandinavica*, **64**, 415-422.
- CANOUNE, H. L., & LEYHE, E. W. (1985). Human versus computer interviewing. *Journal of Personality Assessment*, **49**, 103-106.
- CLARIS (1991). *HyperCard 2.1* [Computer program]. Santa Clara, CA: Claris Corporation.
- KOBAK, K. A., REYNOLDS, W. M., ROSENFELD, R., & GRIEST, J. H. (1990). Development and validation of a computer-administered version of the Hamilton Depression Rating Scale. *Psychological Assessment: A Journal of Consulting & Clinical Psychology*, **2**, 56-63.
- LEVINE, S., ANCILL, R. J., & ROBERTS, A. P. (1989). Assessment of suicide risk by computer-delivered self-rating questionnaire: Preliminary findings. *Acta Psychiatrica Scandinavica*, **80**, 216-220.
- MATARAZZO, J. D. (1986). Computerized clinical psychological test interpretations. *American Psychologist*, **41**, 14-24.
- MUÑOZ, R. F., GONZÁLEZ, G. M., & STARKWEATHER, J. (1991, August). *Bilingual automated screening for depression using computerized speech recognition*. Paper presented at the 99th Annual Convention of the American Psychological Association, San Francisco.
- NOYES, J. M., HAIGH, R., & STARR, A. F. (1989). Automatic speech recognition for disabled people. *Applied Ergonomics*, **20**, 293-298.
- PÉREZ-STABLE, E. J., MIRANDA, J., MUÑOZ, R. F., & YING, Y. W. (1990). Depression in medical outpatients: Underrecognition and misdiagnosis. *Archives of Internal Medicine*, **150**, 1083-1088.
- RADLOFF, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, **1**, 385-401.
- RICHARDS, J. S., FINE, P. R., WILSON, T. L., & ROGERS, J. T. (1983). A voice-operated method for administering the MMPI. *Journal of Personality Assessment*, **47**, 167-170.
- ROBINS, L. N., HELZER, J. E., ORVASCHEL, H., ANTHONY, J. C., BLAZER, D. G., BURNAM, A., & BURKE, J. D., JR. (1985). In W. W. Eaton & L. G. Kessler (Eds.), *Epidemiological field methods in psychiatry: NIMH Epidemiological Catchment Area program* (chap. 8). New York: Academic Press.
- ROSEN, L. D., SEARS, D. C., & WEIL, M. W. (1987). Computerphobia. *Behavior Research Methods, Instruments, & Computers*, **19**, 167-179.
- SELM, P. A., KLEIN, M. H., GRIEST, J. H., SORRELL, S. P., & ERDMAN, H. P. (1990). Computer-administered cognitive-behavioral therapy for depression. *American Journal of Psychiatry*, **147**, 51-56.
- SHAPIRO, S., SKINNER, E. A., KESSLER, L. G., VON KORFF, M., GERMAN, P. S., TISCHLER, G. L., LEAF, P. J., BENHAM, L., COTTLER, L., & REGIER, D. A. (1984). Utilization of health and mental health services: Three Epidemiological Catchment Area sites. *Archives of General Psychiatry*, **41**, 971-978.
- TRONCONI, A., BILLI, M., BOSCARIELI, A., GRAZIANI, P., & SUSINI, C. (1989, October). *Graphics with special interfaces with disabled people*. Paper presented at the Biennial Conference on Augmentative and Alternative Communication, Anaheim.
- WEISSMAN, M. M., BRUCE, M. L., LEAF, P. J., FLORIO, L. P., & HOLZER, C. (1991). Affective disorders. In L. N. Robins & D. A. Regier (Eds.), *Psychiatric disorders in America: The Epidemiological Catchment Area study* (pp. 53-79). New York: Free Press.
- WEISSMAN, M. M., SHOLOMSKAS, D., POTTINGER, M., PRUSOFF, B. A., & LOCKE, B. Z. (1977). Assessing depressive symptoms in five psychiatric populations: A validation study. *American Journal of Epidemiology*, **106**, 203-214.
- WILSON, F. R., GENCO, K. T., & YAGER, G. G. (1985). Assessing the equivalence of paper-and-pencil vs. computerized tests: Demonstration of a promising technology. *Computers in Human Behavior*, **1**, 265-275.