Human-Computer Interaction: 
Conducting a Validation Usability Test

by

Edward A. Debes, II

Submitted in partial fulfillment of the requirements for
DCTE 820

Graduate School of Computer and Information Sciences
Nova Southeastern University

2002
Table of Contents

List of Figures and Tables  iv

Chapters

1. Introduction and Statement of Problem  1
   Test Population  6
   Test Environment and Equipment  6
   Researcher or Monitor Role  6
   Background and Description of the Product  7
   Ease of Use and Ease of Learning  8
   Statement of Informed Consent  8
   Barriers 8

2. Review of Literature  8

3. Methodology  13
   Measurements  19
   Usability Metrics  19
   Test Design  20
   Instructions Related to Usability Evaluation  20
   Detailed Description of the Test Procedure and Instructions to the Participants  21
   A Description of the Task List  21

4. Results  18
   Observations-Qualitative Metric  24
   Observations Relating to all Test Subjects  24
   Results-Quantitative Metric  27
   Errors  28

5. Recommendations and Conclusion  27

Reference List  33

Appendixes
A. Statement of Informed Consent  39
B. Human Computer Interaction Usability Test Subject Questionnaire  41
C. Software Evaluation Questionnaire  43
D. User One Dictation Results  45
E. User Two Dictation Results  46
F. User Three Dictation Results  47
G. User Four Dictation Results   48
H. User Five Dictation Results   49
I.  List of Comments Made by Test Subjects   50
J. Task List Distributed to Users for Dragon Naturally Speaking   51
List of Tables

Tables

1. Example A: Abbreviated Task List 23
2. Samples of Spoken Narrative Used in Dictation Exercises 25
3. Time Taken to Complete Exercises by User 27
4. Errors Committed by Users During Each Exercise 28
5. Errors in Transcription of Samples 29
Chapter 1

Introduction and Statement of Problem

Recent studies indicate employers are not satisfied with the verbal communication skills of recent college graduates who are entering the work force. The Metro Atlanta Chamber of Commerce conducted a survey in 1999. Ninety-percent of the companies responding to this survey indicated that many of today's college graduates lack the basic speaking and listening skills most desired by employers (Joyner, 1999).

The Pennsylvania State System of Higher Education (PSSHE), which consists of fourteen, four-year educational institutions conducted a similar survey of employers in 2000 (Domin, 2001). Results of the Year 2000 PSSHE survey parallel those of the Metro Atlanta Chamber of Commerce study. Recent graduates of PSSHE institutions have fallen short of employer's expectations (Domin, 2001).

Generalizing the data from this study indicates graduates of PSSHE institutions lack the ability to communicate (speaking and listening) on an interpersonal and small group level, both casually with fellow employees and professionally with subordinates, superiors, and customers (Domin, 2001). Both the Metro Atlanta Chamber of Commerce and the PSSHE surveys reveal that college graduates lack the verbal communication skills required to enhance job performance and the performance of co-workers.
Employers do not feel that institutions of higher learning are graduating students possessing adequate communications skills. In order to address this criticism, colleges and universities must improve their verbal communications curriculums. Therefore, colleges and universities need to consider other solutions to improve these communication skills while staying within current constraints of faculty, facilities, budget, and curricula (Hobgood, 2000).

The development of workplace communication skills begins inside the classroom (Cronin & Grice, 1991). Tutors and coaches are available at most colleges and universities to assist students outside class to improve communication skills (Cronin & Grice, 1991). However, the current paradigm does not leave students with enough time to actually practice and develop verbal communications skills.

Computing technology may offer a means by which undergraduates can individually improve their verbal communications skills, specifically through the integration of speech recognition software into communications curriculums. Speech recognition software could provide students with an interactive way to practice speech presentations outside the classroom. A student dictates an assigned speech to a personal computer via a microphone. The speech recognition software “captures” or records the speech presentation in a digital form as the student is dictating. When the student completes presenting his or her speech, the software program allows the student to print a copy of the dictated exercise to review. The visual learner is able to read a review of the speech presentation.

This type of multimedia, computer-mediated feedback provides a way for both visual and aural learners to assess their work (Dringus, 1999), improve deficiencies, and eventually become effective verbal communicators. The speech student can use the speech recognition program to
provide a visual record (transcript) for review of his or her performance, consequentially reducing future errors and eliminating filled pauses.

Speech recognition programs are human-computer interactive. Observing a human testing a software program gives evaluators insight into technological problems and barriers, they may never witness otherwise. Testing speech recognition products for universal usability is an important step before considering the product to be a viable solution for improving undergraduate students’ communication skills. In other words, speech recognition software must be easy to use for most, if not all, undergraduates.

Vanderheiden (2000) defines universal usability as a focus on designing products so that they are usable by the widest range of people operating in the widest range of situations as is commercially practical. Usability testing helps determine if the final software product meets the goals of the software product designers. In addition, usability testing may uncover product goals not originally defined by the designers. The information provided by usability testing can lead to improving product design and correcting user interface and design problems. Such information enables consumers to make educated comparisons of similar software products. Usability evaluation helps determine if a product’s technology is an obstacle, rather than an aid, to the user while he or she performs prescribed tasks.

Most software products require usability testing. Usability testing is necessary because designers and developers do not always make correct decisions while building a software product (Wilson & Coyne, 2001). The design principles and guidelines for software development are not always sufficient, and informal feedback is inadequate for evaluations of the user interface (Mandel, 1997).
The researcher will explore the usability evaluation process, as it relates to speech recognition software. Dragon Naturally Speaking Version 6 will be the product evaluated. The goals of this speech recognition software are to do the following: recognize the human voice; transcribe words spoken by the human voice; and convert transcribed words to text as part of a word-processed document. Through this evaluation, the researcher anticipates that the results will demonstrate Dragon Naturally Speaking Version 6 is a software product that does not allow the computing technology to interfere with the ability to concentrate on building their human communication skills (Appleton, 1993).

In usability evaluations, problems include, but are not limited to the following:

Do novice users inadvertently access advanced features?
- Do the screens present information in a manner consistent with previous screens?
- Is the response time a problem?
- How accessible and helpful are the Help functions?
- What obstacles prevent the completion of a task?
- Can the user perform a specified task within established time constraints?
- If the product is a subsequent version of a product, is the new release more usable?

In the usability test that the researcher will conduct for Dragon Naturally Speaking Version 6, the primary concern is ease of use. The researcher has been using the product for approximately ten months as a tool for dictating documents, memos, exam, letters, email, and doctoral research papers. Problem statements for this evaluation include, but are not limited to the following:

How easy is it to set-up a new user in Dragon Naturally Speaking Version 6?
- How difficult is it for the user to navigate the software to complete the enrollment phase?
- How accurate can the software transcribe the spoken word and convert it to text in a document?
- If a user makes an error, how difficult is it for the user to recover so he or she can resume using the software?

According to Rubin (1994), the overall goal of usability testing is to identify and correct usability deficiencies. Usability testing insures that the product is easy to learn, easy to use,
satisfying to use, and provides the utility and functionality highly valued by the population for which the product is designed (Rubin, 1994; Vanderheiden, 2000). Rubin states that without specific goals or objectives usability testing may become too informal and unfocused. There are four types of usability tests: exploratory, assessment, validation, and comparison. The purpose of the researcher’s study will be to conduct a validation usability test. A validation usability test confirms content and certifies a product’s usability (Rubin, 1994).

Product leaders and developers often cannot adequately define what the product is supposed to do for the user (Wilson and Coyne, 2001). Product leaders and developers make defining statements, which are vague and lack direction. For example, statements such as, “The product should be intuitive and easy to learn”, are not clear enough to serve as test goals and objectives (Mandel, 1997).

As a result, the usability-testing researcher must often define his or her testing goals. The following is a list of the goals and objectives that the user will attempt to accomplish while testing the Dragon Naturally Speaking Version 6 software product.

Start Dragon Naturally Speaking Version 6 software application.
Setup microphone.
Complete enrollment phase.
Dictate three statements and convert statements to text.
Close the program.

According to Shneiderman (1998), it is important to write a clearly defined task list. This list details the planned testing activities. The author will measure test goals and objectives by taking note of the errors and difficulties test subjects encounter while operating Dragon Naturally Speaking Version 6 as a new user.

The Methodology section of this paper presents an abbreviated task list.
**Test Population**

A wide range of users can use Dragon Naturally Speaking Version 6. The researchers will chose a test population representing the wide range of users that can use Dragon Naturally Speaking Version 6. The researcher will select five adults for this usability test. The test subjects represent each of four age groups: 15 to 24, 25 to 35, 36 to 45, and 46 to 60. All the test subjects have PC computer experience and have experience using a word processor. The test subjects are the researcher’s friends and neighbors. All test subjects are over 18 years of age. All the test subjects have had no prior experience with speech recognition software.

**Test Environment and Equipment**

The testing location will be the basement of the researcher’s home. The basement is finished as a recreation room and furnished in a modern design. Distracting street and domestic noises (i.e. telephone, doorbells, and family pets, etc.) are limited by the basement’s insulated location. The testing computer is an IBM PC clone equipped with a 1-gigahertz processor and 512 megabyte of random access memory. The microphone used for this test is the Telex H 531 USB digital model. Before the test, the researcher will install the Dragon Naturally Speaking Version 6 software product on the PC.

**Researcher or Monitor Role**

The researcher will act as an observer. His role will be to record, study, and watch the test subjects. The researcher is already acquainted with how to use the Dragon Naturally Speaking Version 6 software product. Near the end of the development stage and prior to releasing the software product for distribution, the software designer who created the product may lack the
objectivity and empathy needed to understand the challenges a new user faces to operate the product successfully. As an observer, the researcher is aware that he has the experience to understand and recognize the learning challenges of the test subjects more objectively than the software designer. Feedback based on observing users frustrations with learning a new software product firsthand is very valuable. Such feedback can do more to convince a designer to revise a product than weeks of discussions and negotiations (Rubin, 1994).

**Background and Description of the Product**

ScanSoft, Incorporated, manufactures Dragon Naturally Speaking Version 6 (Refer to Figure 1 for an example of the packaging). ScanSoft is the leading supplier of imaging, speech, and language solutions computing technology. ScanSoft’s products automate a wide range of manual processes resulting in saving time, increased worker productivity, and improved customer service.

Dragon Naturally Speaking Version 6 is the market-leading desktop dictation solution for the PC, allowing people to use their voice to create documents at up to 160 words-per-minute. Integrated with Microsoft® Office and Corel® WordPerfect®, Dragon Naturally Speaking Version 6 allows the computer to become a powerful voice-driven tool that can recognize not just what the user wants to type, but what the user wants to do. Designed for American English, the software product is adaptable for Australian, British, Indian, and Southeast Asian speakers of English.

It is the author’s opinion that as student interest in new and innovative ways to develop speaking skill increases, and the price of purchasing PC’s and software continues decreases; new and innovative ways to improve speaking skills will rise. In the author’s opinion, the delivery of
speech improvement materials in an online environment will become a valuable and practical educational solution.

Figure 1 – Dragon Naturally Speaking Version 6 packaging.

Ease of Use and Ease of Learning

The researcher understands that the Dragon Naturally Speaking Version 6 software has a wide audience of users. Therefore, addressing usability issues is important. The researcher believes the software is usable and easy to learn and use out of the box. This is especially true for those individuals who have an interest in the tools and features the product offers.

Statement of Informed Consent

A signed Statement of Informed Consent is required of all test subjects. The researcher will use a Statement of Informed Consent form based on the Shneiderman design (Shneiderman, 1998). See Appendix A for a copy of the form used by the researcher.

Barriers

Barriers to conducting the usability test include, but are not limited to the following:

The test population’s ability to speak well and properly annunciates the words.
The reading ability of the test population.
The physiological condition of the test population (colds, allergies), which may affect the
ability to speak.
The test population’s understanding of basic Windows personal computer commands.
A definition of speech recognition is the ability of the personal computer to recognize natural human speech (Newman, 2000). Using speech recognition software implies a new frontier in learning and requires a pioneering attitude (Abramson, 2000; Bruce, 2001; and Harris, 2001). Increased processing speeds, inexpensive storage, and better sound reproduction and recording quality are just a few of the reasons speech recognition is increasing in popularity (Newman, 2000). Speech recognition software is becoming easier to use and more accurate in its ability to identify words correctly (Newman, 2000).

Bork (2000) is a pioneer in the area of speech recognition. He has identified speech recognition as a tool to use with new learning paradigms based on the classical design of one-on-one teacher-student tutorial learning. Bork argues that the personal computer, as a tutor, will never approach the quality of a highly skilled human tutor, but Bork considers personal computer-mediated learning paradigms to be a leap forward from the information-transfer paradigm currently being used. The new personal computer-mediated learning paradigm relies on student-centered interaction at a high level and involves multimedia interfaces such as microphones and sound interface cards (Bork, 2000).

Shneiderman (1998) identifies four variations for speech technology: speech store and forward; speech generation; discrete-word recognition; and continuous-speech recognition. Commonly used in fee-based telephony, *speech store and forward* is the process of storing a
single voice message and forwarding the message to one or more subscribers. Those subscribers can play and replay messages; respond to the original message; delete or archive messages. An example of speech store and forward is the voice mail systems of large phone systems or wireless technology.

Speech generation or speech synthesis was created by software designers in order to “humanize” computers, according to Shneiderman (1998) However, one of the common problems associated with speech synthesis is that the novelty of a computer talking decreases with the passage of time. These disembodied computer generated voices becomes an annoyance, rather than a help to consumers. Examples of speech synthesis include the soft-drink vending machine that politely says “thank you” after you pay for and make your drink selection. Another example is automobiles that remind you that “your door is ajar” when the door is left open after a specified amount of time. A final example is secure area warnings that sternly shout “danger, danger” when someone nears a restricted or protected area.

By contrast, discrete-word recognition is based on the recognition of individual words spoken by a specific person with 90 to 98 percent reliability. The basis of this assumption is the existence of a small 20 to 200-word vocabulary and only after the user has completed a vocabulary training session. A common application of discrete word recognition is voicing command and control instructions to a computer. Examples of discrete-word recognition include instructing the personal computer to start a program, open a file, and save a file. Continuous-speech recognition allows the user to dictate text to a PC. This text can be included in the saved document file for later review and refinement.

According to Newman (2000), current speech recognition programs enable users to accomplish two broad operation tasks. The first task is the dictation of text into a personal
computer (continuous-speech recognition). The second task is voicing command and control
instructions to a personal computer or controlling a personal computer by speaking commands
instead of typing them (discrete-word recognition). While most of the speech recognition
software products available today focus on command and control, many of these products also
allow the user to dictate to a personal computer, as well (Newman, 2000).

The most widely used application of voice recognition by business and industry is dictation
(Weinschenk & Barker, 2000). According to Newman (2000), the dictation program *fools* the
personal computer into thinking a person is actually typing. The software operates by actually
placing words in a text file. Found mostly in commercial businesses, legal firms and physician's
offices, voice recognition software helps reduce the amount of work hours associated with
human voice to text translation (Weinschenk & Barker, 2000).

According to Newman (2000), this process of converting speech to text is not as
straightforward as assumed. The user must train the software to recognize the user's voice and
the user’s annunciation style. This level of human-computer interaction according to Newman is
the *enrollment* period. During the enrollment period, the user reads text into the personal
computer via a microphone located in a controlled environment. Also known as *speaker-
independent training* (Shneiderman, 1998), the speech recognition software learns the user’s
inflections, delivery, and intensity levels (Weinschenk & Barker, 2000). Without the enrollment
process, the software may make undesirable assumptions about the user (Newman, 2000).

Command and control voice recognition was the first successfully developed consumer
application, largely because the product uses a limited vocabulary (Newman, 2000). Voice
command sequences spoken in order, such as "file, open, project budget," will open a Microsoft
process of giving command and control instructions to a personal computer is *discrete dictation*. Newman and Shneiderman observe that *discrete dictation* is an ideal way for handicapped individuals to operate personal computers, especially those who have limited hand control or do not have hands. Vanderheiden (2000) concurs with Newman and Shneiderman that discrete dictation provides a method with which handicapped individuals can access personal computer, but also advances the goal of universal usability or as designing products so they are usable by a wide range of people operating in the widest range of situations as is commercially practical.

According to Newman (2000), computer speech recognition is the result of integrating several software development approaches into a system that enable computers to recognize natural human speech. Rubin (1994) writes that the development of a system or software product is an attempt to improve human performance in some capacity. Since software is necessary for a computer to function and humans interact with computers, system and product designers need to consider usability issues far more than they have. As computing technology prices decrease, the amount of possible users increases and as result the number of computer savvy users decreases. According to Wilson & Coyne (2001), software developers believe it is easier for humans to adapt to a product, than to develop a product that adapts itself to more than one user. Many developers have chosen the “human must adapt” methodology in product development. As a result, software developers, who are generally engineers, are better able to work with concrete scientific issues than with human related issues (Wilson and Coyne, 2001). In the past, when end users were computer savvy, this was not as much of a problem.

Many software products are not simply standalone applications with easy to follow instructions. Software products are often networked, dynamically linked, Web-enhanced sophisticated applications for users with a wide range of skills and abilities. With such
complexity confronting users, how do users know whether these products actually deliver what they advertise? Does the design of the product address usability issues during the product development? According to Mandel (1997), only usability testing can provide the answers to these questions.

Usability testing methods include the think aloud method, heuristic evaluation, cognitive walkthrough, expert reviews, guidelines review, consistency inspection, coaching method, observation, focus groups, usability inspection methods, and the use of interviews and surveys. Usability inspection has become one of the most widely used methods. In usability inspection, evaluators inspect usability aspects of a product. The defining characteristic of this method is the reliance on the evaluator’s judgment as a source of feedback (Jacobsen, Hertzum, & John, 1998; Nielsen, 1995).

In heuristic evaluation, the user interface evaluator studies the product and looks for properties that lead to usability problems. The evaluator compares the product’s compliance with generally known or recognized usability principles. Evaluators with experience are likely to find many of the problems with software usability to be the result of user interface design. However, non-experts can also find user interface problems. When considering heuristic evaluation, which includes general costs, the user interface design, and the testing process, three to five evaluators seem to work best for this method (Leceroif and Paterno, 1998).

Kubie and Melkus (1996) write that there are three major approaches to usability evaluation. These are heuristic reviews, walkthroughs, and laboratory tests. In walkthroughs, users or a project team review all or part of a software application and make judgments after seeing a demonstration of the product. In a laboratory test, users make judgments based upon hands-on experience with the product for a defined set of tasks.
Thinking aloud is a widely used method for usability evaluation. However, the think aloud method is a controversial method. Often the think aloud method is used in conjunction with other evaluation techniques. In the think aloud method, tests are set up in which a user is given a number of realistic tasks to perform. The user works through these tasks talking aloud, explaining what he or she is thinking to the evaluator (Waes, 2000).

As previously mentioned, Shneiderman (1998) concurs with Rubin (1994) that the overall goal of usability testing is to identify and correct usability deficiencies. The result of usability testing should insure that the product is easy to learn and use, is satisfying to use, and provides utility and functionality. These design attributes are highly valued by the population using the software product (Sullivan and Matson, 2000; Vanderheiden, 2000).

Most usability scholars believe current usability design principles and guidelines are not sufficient. The design process requires more usability information. Mandel (1997) and Wilson & Coyne (2001) believe a gap exists between the end-product vision of the designer and what users really want and expect from the end-product. People differ and because people differ there is no “average” user that designers can use as a model in the development of a product. Wilson and Coyne support usability testing early and throughout the design process. A user interacting with the product is a reliable way to collect information (Shneiderman, 1998).

Usability tests also result in providing both quantitative and qualitative data. Quantitative reliability is a result of statistical significance (Terrell, 2001). For a usability evaluation to be reliable the evaluator must be able to determine what the results from the sample used indicates about the whole population. Inferential statistics accomplishes this task (Terrell, 2001).

An external consultant can review the findings to improve qualitative reliability. According to Gay and Airasian (2000), this will provide a safeguard against the biases of single evaluators.
Obtaining the user’s direct impressions, observing the user’s actions, and verifying the conclusions or observations of the user can lead to qualitative internal validity (Gay & Airasian, 2000).

The selection of participants, based on the make-up of the end-user population helps establish qualitative external validity (Gay & Airasian, 2000). Using realistic tasks, the equipment typical of the end-user and realistic prototypes of documentation supplied with the product is a further requirement to establish validity. Using the think-aloud method grounds the qualitative data as directly as possible in user perceptions and experience (Hughes, 1999). The researcher plans to use the think aloud method for this usability test. A byproduct of usability tests is collecting quantitative and qualitative data. Such data can provide advantages to software companies who consider the data as part of their development process over companies who do not perform usability testing.

According to Rubin (1994), a usability test typically has four stages, preparation, introduction, the test, and debriefing. In preparation for the experiment, the researcher needs to do the following: make ready the testing room; make the computer system start-state ready for the test; and, make ready the test materials, instructions, and questionnaires for distribution to the test subjects (Rubin, 1994).

During the introduction, the researcher welcomes the test subject and provides a brief explanation of the test’s purpose. The researcher then discusses the test procedure with the test subjects. Rubin (1994) suggests incorporating the following statements into the description of the test procedure:

- The purpose of this test is to evaluate the software, not the user.
- You are encouraged to speak freely during the test.
- The test results will improve the product.
- The test results are confidential.
Participation in the test is voluntary and you (the user) may stop the test at anytime. You are entitled to notification of and an explanation of any video or audio recording that may take place during the test. You are encouraged to ask questions, since it is an objective of the test to determine what is unclear. You are entitled to any specific instructions, such as an explanation of the think aloud procedure. Do you have any questions before beginning the test?

During the test, the researcher should normally refrain from interacting with the user. Even if the user gets into severe difficulties, the researcher should refrain from providing assistance. If the user is clearly having difficulty performing the task required by the test (Nielson, 1993), the researcher can provide assistance as an exception.

A designer can gain information from observing a user struggle and become frustrated when working with a product. Actually seeing this struggle first hand will do more to convince a designer to continue the efforts to improve a product’s usability (Rubin, 1994).

Neilson (1993) and Rubin (1994) examine quantifiable usability measurements that include, but are not limited to the following:

The time it takes users to complete a task.
The number of tasks a user completes within the specified time limit.
The ratio of successful user interactions to user errors.
The time users spent recovering from errors.
The number of user errors.
The number of unused commands and features.
The number of user command and features.
The number of features the user can remember after the test.
The frequency of use of manuals and help features during the test.
How many times using manuals and help features solved the user’s problem(s).
The proportion of user statements made during the test that were positive versus negative toward the system.
The number of times there was clear frustration or joy during the test.
The proportion of users who said they prefer using the system over that of a specified competitor.
The number of times the user had to work with an unsolvable problem.
The amount of user “dead time” during the test.
The number of times distraction occurred during the test.
A scheduled debriefing session occurs at the completion of the test. Questionnaires and interviews are useful methods for collecting data on how users utilize computer systems. Questionnaires are also useful for determining what features users particularly like or dislike about the software product. The researcher should have the questionnaires completed before any other discussion ensues (Nielsen, 1993).
Chapter 3
Methodology

Measurements

During the usability evaluation of Dragon Naturally Speaking Version 6, the researcher will obtain the following measurements:

- The time it takes the user to complete the entire task list.
- The time it takes the user to complete each step in the task list.
- The number of errors committed by each user while navigating the software to complete the task list.
- The number or errors resulting from the software incorrectly transcribing the dictation of drill exercises.
- A list of the users’ positive and negative comments about their experiences with the software during the usability evaluation test.

The researcher will provide a questionnaire to the user at the conclusion of the evaluation. The purpose of the questionnaire is to measure user satisfaction with Dragon Naturally Speaking Version 6 (Kirakowski, 1994; Nielson, 1993; Shneiderman, 1998). Once the questionnaire is completed, the researcher will enter into a discussion with the user for the purpose of clarifying and gaining more information about the user's experience.

Usability Metrics

The qualitative usability metric is in the form of the researcher’s observations of the user. Simply, the researcher will take note of each user's spoken words and actions and record these
observations on paper. Measurement of quantitative metrics includes but is not limited to the following (Hughes, 1999; Mayhew, 1999):

- The time it takes each user to complete the overall task.
- The time it takes each user to complete each exercise.
- The amount of errors encountered by each user when navigating the software.
- The number of positive and negative comments made by each user.
- The number of tasks completed successfully by each user.
- The number of times the user was frustrated.

**Test Design**

The users selected for this test are a small representative sample of this product's buying population. Each participant will work independently. The testing process consists of a five-day period, with one user a day tested by the researcher. Each participant will work through a task list, which consists of a series of exercises designed to train the software to recognize the user’s voice. A series of steps makes up each exercise. These exercises are typical of what first time users would ordinarily encounter when using the software. The exercises consist of the following:

- Setup a new user.
  - Complete the enrollment/training phase.
  - Dictate three narrative samples and compare results with the original.

**Instructions Related to Usability Evaluation**

Each user in this study will work under the premise that the software is a productivity tool. Each user in this study will be encouraged to use the product "as is”. The product ships with a software license form, the program on CD-ROM, and a software manual. The researcher will create a task list that will measure ease of product use in an orderly fashion.
Detailed Description of the Test Procedure and Instructions to the Participants

Each participant or user is deserving of respect. In addition, users must understand that it is the software under examination and not the user (Shneiderman, 1998). Each user will be welcomed into the home of the researcher. The researcher will discuss the software product.

Next, the researcher will request each user to complete and sign the Statement of Informed Consent (see Appendix A). The researcher will provide answers to any user having questions or concerns at this point in the process. The researcher will physically hand to the user a task list as presented in Appendix J.

After that, the researcher will notify each user that the researcher will not be available to provide assistance during the evaluation. The researcher will inform each user that there is a time limit of 30 minutes to complete the entire task list. The researcher will take notes, keep track of errors, encourage the talk aloud method, and try not to distract the user.

Upon completion of the test, the researcher requests that users complete a questionnaire and survey (see Appendix B and C). Next, the user and researcher will participate in a posttest interview. If needed, during the debriefing, the researcher will show the user how to navigate the software, how to use the software as a productivity tool, and answer any remaining questions.

A Description of the Task List

Usability scholars concur that test tasks need to be as representative as possible to actual practice. Tasks must provide reasonable coverage of the most important uses of the product. The tasks need to be small enough to complete within the allotted time. However, the tasks need to be not so small as to be trivial.
The user should expect a task list specifying what expectations the researched expects he or she to meet. In addition, the researcher presents tasks in a written form, i.e. typed, or word-processed. Users are encouraged to ask questions about the task list before beginning the evaluation in order to help eliminate any confusion regarding either the task list or the test itself. The researcher's presents the test tasks in a business-like manner. Test tasks structured using the above guidelines will increase the user's confidence (Nielsen, 1993).

**Task List (Example A)**

Table 1 presents an abbreviated list of tasks for the usability evaluation of Dragon Naturally Speaking Version 6. Refer to Appendix J for an actual copy of the task list distributed to test subjects.
**Table 1 – Example A: Abbreviated Task List.**

<table>
<thead>
<tr>
<th>Step Number</th>
<th>Abbreviated Task List for Usability Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step One</td>
<td>Start program by clicking on &quot;Dragon&quot; icon.</td>
</tr>
<tr>
<td>Step Two</td>
<td>Select “Setup New User”.</td>
</tr>
<tr>
<td>Step Three</td>
<td>Complete required information fields.</td>
</tr>
<tr>
<td>Step Four</td>
<td>Place microphone headset on head.</td>
</tr>
<tr>
<td>Step Five</td>
<td>Test microphone record level.</td>
</tr>
<tr>
<td>Step Six</td>
<td>Complete voice testing phase of microphone.</td>
</tr>
<tr>
<td>Step Seven</td>
<td>Select “Start” to begin training enrollment process.</td>
</tr>
<tr>
<td>Step Eight</td>
<td>Read each panel presented until told process is completed.</td>
</tr>
<tr>
<td>Step Nine</td>
<td>Select “Dictate”.</td>
</tr>
<tr>
<td>Step Ten</td>
<td>Start Microsoft Word by selecting icon on desktop”.</td>
</tr>
<tr>
<td>Step Eleven</td>
<td>Activated microphone.</td>
</tr>
<tr>
<td>Step Twelve</td>
<td>Begin dictation by reading copy provided by researcher.</td>
</tr>
<tr>
<td>Step Thirteen</td>
<td>Close microphone.</td>
</tr>
<tr>
<td>Step Fourteen</td>
<td>Save work under “Test User -Name”.</td>
</tr>
<tr>
<td>Step Fifteen</td>
<td>Close Microsoft Word.</td>
</tr>
<tr>
<td>Step Sixteen</td>
<td>Close Dragon Naturally Speaking.</td>
</tr>
</tbody>
</table>
Chapter 4

Results

**Observations-Qualitative Metric**

The researcher provided each user with specific directions before starting the test. The researcher indicated he was completing an academic assignment. He informed each user of his or her role in the assignment. Next, the researcher introduced the user to the product title, the manufacturer's name, and explained why he chose the product for this test. The researcher again encouraged the user to think aloud during the test. The researcher asked if there were any questions and answered such questions if they occurred. After receiving the task list (Appendix J), the researched asked each user to begin the test.

**Observations Relating to all Test Subjects**

Test subjects performed all tasks with minimal errors. All test subjects cooperated with the researcher during the test and all completed the test with the specified time limit. All subjects successfully completed the enrollment/training process and were able to dictate text to the computer.
Table 2 – Samples of Spoken Narrative Used in Dictation Exercises.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Examples of Spoken Narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample A</td>
<td>I am now able to talk to my computer!</td>
</tr>
<tr>
<td>Sample B</td>
<td>Dear Britt: How was Germany? I am so glad you brought me chocolate! We will share the treats at your party. Love, Angie.</td>
</tr>
<tr>
<td>Sample C</td>
<td>Good morning, Boys and Girls! Today we will have music and we will be reading a story about a bear. How many of you brought a bear with you today? You may get out your bear and share it with the person next to you. Tell your partner your bear’s name and how you go this special bear.</td>
</tr>
</tbody>
</table>

All users began the test by reading the first page of the testing materials prepared by the researcher. Following the directions provided on the first page, each user launched the application and began the “setup new user” feature. Each user followed the prompts and answered all questions relating to establishing a new user.

Once the user completed the new user setup, the user began the training/enrollment phase. As each user progressed through the test, he or she relied less on the written materials and more on the user interface. This was a positive indicator. The software allowed the user to focus on the current task, instead of diverting attention to reading supporting materials or solving problems unrelated to the current task.

The enrollment phase trains the computer to recognize the user's voice. The full enrollment program involves reading eight different documents, containing at least fourteen paragraphs each. Completing the entire process can take a few hours. The product designers suggest that the user complete enrollment over a number of days and not over a few hours. In the interest of time, test subjects were only required to complete the first level of training.
The first level of training requires each user to read sixteen short paragraphs of information. This information includes instructions on how to train the software to recognize the user’s voice. In addition, this introductory information educates the user on the science of speech recognition.

Dragon Naturally Speaking Version 6 software reads and analyzes all the training paragraphs dictated to the program. The software automatically places the user in dictation mode. The software requests that the user launch a word processing program, activate the microphone, and begin to dictate text. For this study, the researcher used the Microsoft word processing program, Microsoft Word.

The researcher provided three samples for dictation. Refer to Table 2 for an example of these samples. The software prompted each user to read each sample three times in succession with a slight pause in-between each reading. As each user read a sample, the PC and software application recognized the user’s voice, transcribed the voice input to text, and placed the text into a blank Word file. Refer to Appendixes D through H for the results of these dictations. Next, the software prompted each user to save the text file they had created and then close the Word and Dragon Naturally Speaking Version 6 applications.

As previously mentioned, the researcher requested each user to complete a questionnaire and participate in a debriefing session. Refer to the Results - Quantitative-Metric section of this paper for a review of these questionnaires. Refer to Appendix I for a list of quotes recorded during the debriefing sessions.

The following exceptions are noted. User Two took the longest amount of time to complete the entire exercise. User Two took longer to dictate the exercises than the other users. Although User Two followed the directions correctly, he did not activate the Word document. As a result, the PC did not transcribe any of the words spoken by the user. The user continued to read all
three samples, three times each. The user never looked at the screen until the final reading was finished. When the user looked at the screen, he noticed there was no text visible. The error took him about 3 minutes to solve. Once he solved the error, he was able to dictate the samples provided by the researcher.

Results-Quantitative Metric

The usability test conducted by the researcher measured the following: how easily and how fast a user can complete the prescribed tasks; how many errors a user committed while using the software; and how accurate is the rate of the spoken word verses transcribed word. An important result of the usability test is that the pace chosen by the users allowed them to successfully navigate the user interface within the 30-minute period. The data in Table 3 indicates that students performed the prescribed work within the 30-minute period.

Table 3 – Time Taken to Complete Exercises by User.

<table>
<thead>
<tr>
<th>Exercise Tasks</th>
<th>User 1</th>
<th>User 2</th>
<th>User 3</th>
<th>User 4</th>
<th>User 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to Complete First Level of Enrollment and/Training</td>
<td>8:52</td>
<td>10:57</td>
<td>7:06</td>
<td>8:11</td>
<td>8:02</td>
</tr>
<tr>
<td>Time to Setup Software and Dictate Three Samples, Three Times</td>
<td>7:10</td>
<td>13:30</td>
<td>5:56</td>
<td>5:30</td>
<td>6:54</td>
</tr>
<tr>
<td><strong>Total Time</strong></td>
<td><strong>19:22</strong></td>
<td><strong>28:08</strong></td>
<td><strong>15:18</strong></td>
<td><strong>18:23</strong></td>
<td><strong>19:35</strong></td>
</tr>
</tbody>
</table>
Errors

Table 4 presents the amount of errors committed by each user during each exercise as observed by the researcher.

**Table 4 – Errors Committed by Users During Each Exercise.**

<table>
<thead>
<tr>
<th>Description of Exercise</th>
<th>User 1</th>
<th>User 2</th>
<th>User 3</th>
<th>User 4</th>
<th>User 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete New User Setup</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Complete Enrollment/Training</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dictate Three Samples, Three Times</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total Errors</strong></td>
<td><strong>8</strong></td>
<td><strong>9</strong></td>
<td><strong>10</strong></td>
<td><strong>8</strong></td>
<td><strong>6</strong></td>
</tr>
</tbody>
</table>

Table 5 depicts the accuracy rate of each sample. Each user read the sample three times and the software product recorded the results of the reading. The researcher compared the original text with the transcribed text and calculated the amount of errors. Words, sentences, and punctuation marks make up each sample. Punctuation is not a concern. Saying aloud what type of punctuation is included in the text is not normal practice. The listener assumes punctuation through pauses and voice inflections (DeVito, 2002).

Each sample consists of an amount of words or variables: Sample One has nine variables, Sample Two has 23 variables, and Sample Three has 58 variables. This focus of this study is to measure the ability of the software product to improve verbal communication skills. Therefore, accurate transcription of words and sentences is the most important variable.
Table 5 – Errors in Transcription of Samples.

<table>
<thead>
<tr>
<th>Transcription Errors by Sample</th>
<th>User 1</th>
<th>User 2</th>
<th>User 3</th>
<th>User 4</th>
<th>User 5</th>
<th>Accuracy Rate for Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample One (9 variables)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcription Errors Reading 1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>93</td>
</tr>
<tr>
<td>Transcription Errors Reading 2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>91</td>
</tr>
<tr>
<td>Transcription Errors Reading 3</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>93</td>
</tr>
<tr>
<td>Average Accuracy Rate for User in Percent (%)</td>
<td>96</td>
<td>78</td>
<td>93</td>
<td>96</td>
<td>100</td>
<td>93</td>
</tr>
<tr>
<td>Sample Two (23 variables)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcription Errors Reading 1</td>
<td>3</td>
<td>12</td>
<td>8</td>
<td>10</td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>Transcription Errors Reading 2</td>
<td>6</td>
<td>13</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>Transcription Errors Reading 3</td>
<td>3</td>
<td>10</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>78</td>
</tr>
<tr>
<td>Average Accuracy Rate for User in Percent (%)</td>
<td>83</td>
<td>51</td>
<td>62</td>
<td>77</td>
<td>93</td>
<td>73</td>
</tr>
<tr>
<td>Sample Three (58 variables)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcription Errors Reading 1</td>
<td>9</td>
<td>12</td>
<td>9</td>
<td>7</td>
<td>2</td>
<td>87</td>
</tr>
<tr>
<td>Transcription Errors Reading 2</td>
<td>5</td>
<td>15</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>Transcription Errors Reading 3</td>
<td>1</td>
<td>15</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>90</td>
</tr>
<tr>
<td>Average Accuracy Rate for User in Percent (%)</td>
<td>91</td>
<td>76</td>
<td>89</td>
<td>90</td>
<td>95</td>
<td>88</td>
</tr>
</tbody>
</table>

The designers of Dragon Naturally Speaking Version 6 claim the average accuracy rate of the product is 95% and above. This high rate of accuracy is established when the user completes the entire enrollment process. As previously mentioned, time will not allow each user to complete
the entire enrollment process. However, even with only the first level of the enrollment completed, the product scored some impressive accuracy rates.

Sample One set an average accuracy rate of 93% for the five test users. Sample Two did not fare as well. The voice quality of Users Two and Three may have played a part in the reduction of accuracy for Sample Two. User Two’s voice was "boomy" and lacked “space” in-between words. User Three had speech impediments and rounded off many consonants while speaking. Sample Three scored an accuracy rate of 88% with User Two still scoring well below all users in the group. User Five continually scored very high in all Samples. User Five’s voice is mature, and without any speech impediments. User Five read at a moderate speed.

A design characteristic of Dragon Naturally Speaking Version 6 is the software’s ability to continually improve its accuracy rate each time a user utilizes the product. Table 5 supports this fact. Sample One improved from 91% on the Second Reading to 93 % on the Third Reading. Sample Two improved from 70% on the First Reading to 78% on the Third Reading. Sample Three improved on the First Reading from 87% to 90% on the Third Reading.

Think aloud observations and posttest interviews offer a first person view of the user’s experience using the product. Posttest interview data for this usability study indicate users had fun using Dragon Naturally Speaking Version 6. Some users were expecting the software to operate in one way and were surprised when the software operated in quite a different way.

For example, one user assumed the conversion to text would be instantaneous. From a technological standpoint, the conversion process may take a few seconds, but the ability to do a conversion in fractions of a second is unrealistic. Users were not troubled with the complexity and the look of the user interface. Screen colors were easy on the eyes and the text was easy to read for all users.
Survey data presented in Table 6 also supports the posttest interview data. Users agreed that the basic software operation and access to the enrollment process is straightforward and easy to perform. Three out of five users would recommend the product to others, while two remain undecided. All users indicated the software commands were troublesome at some point, and all users agreed the software was not frustrating to use.
Table 6 - Results of Post-Usability Evaluation Survey.

User: 1 2 3 4 5

This is a software evaluation questionnaire for Dragon Naturally Speaking Version 6.

Directions: Please fill in the leftmost circle if you strongly agree with the statement. Mark the next circle if you agree with the statement. Mark the middle circle if you are undecided. Mark the fourth circle if you disagree, and mark fifth or last circle if you strongly disagree with the statement (Kirakowski, 1994).

<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I would recommend this software to my colleagues.</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2. The instructions and prompts are helpful.</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3. Learning to use this software is difficult.</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4. I sometimes do not know what to do next.</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>5. I enjoyed my session with this software.</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6. I find the help information useful.</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7. It takes too long to learn the software commands.</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>8. Working with this software is satisfying.</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>9. I feel in command of this software when I use it.</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>10. I think this software is inconsistent.</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>11. This software is awkward to use when I want to do something that is not standard.</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>12. I can perform tasks in a straightforward manner using this software.</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13. Using this software is frustrating.</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>14. It is obvious that the software was designed with the user' needs in mind.</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>15. At times, I have felt tense using this software.</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>16. Learning how to use new functions is difficult.</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>17. It is easy to make the software do exactly what you want.</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>18. This software is awkward.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>19. I have to look for assistance most times when I use this software.</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>20. The software has not always done what I was expecting.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
Chapter 5

Recommendations and Conclusion

The designers of Dragon Naturally Speaking Version 6 make few assumptions with respect to their users. In the author's opinion, the software, in its current packaged form, does an admirable job of addressing the usability needs of a wide audience. Vanderheiden (2000) emphatically states that it is necessary to insure designers develop software products with universal usability in mind.

It is the author’s opinion that the Dragon Naturally Speaking Version 6 designers created a generally solid software product that almost anyone can use with success. The software recorded and analyzed each test subject's voice successfully. Afterwards, each user could dictate and the PC transcribed the user’s dictation with relative accuracy.

The voice to text transcription application is a proven feature of Dragon Naturally Speaking Version 6. However, using this software application as a tool to improve speech communication among college students is yet unproven. The researcher believes there is one primary barrier associated with utilizing Dragon Naturally Speaking Version 6 in college speech curriculums.

This barrier is the amount of time a student will need to dedicate to completely training the software in order to achieve a high accuracy rate. As previously mentioned, this study focused on the first level of enrollment and ease of use of the product. However, the product requires six additional levels of enrollment before a user can expect to achieve a projected accuracy rating of 98%. The question for this researcher is, “Will the typical college student commit to completing
the few hours of time needed to train the software before actually using the software for improvement of speech skills?”

The researcher believes undergraduates will stay away from this tool in the current form due to the amount of time required to make the software function at an optimum. The researcher believes as future software releases are developed, the time period needed to complete enrollment will decrease and as a result, the product will become more attractive to students. In the researcher’s opinion, with streamlining of the enrollment process, the software could evolve into an effective tool for improving student’s public communication skills.

A suggestion for further research is to select users who are actually undergraduates needing to improve their speech skills. These selected undergraduates would commit to completing all levels of enrollment training. The test population of undergraduates would complete the first level of enrollment and dictate a selected speech. Next, the test population for undergraduates would complete the enrollment process and dictate the sample speech a second time. A comparison of error and accuracy data after the first reading of the speech and the second reading of the speech might reveal helpful information about the effectiveness of Dragon Naturally Speaking in improving speech skills.

In the opinion of the researcher, improving undergraduates’ communication skills is worth exploring for many reasons. As previously mentioned, employers are of the opinion that college graduates are lacking in communication skills. Current undergraduates have been conditioned by today’s the mass communication culture (Cole, 1999). They are used to reacting and responding in a two dimensional plane.

Employers understand that the key to increased company revenues is successful human interactions focused on a particular goal. How we, as individuals, communicate with other individuals reveals our personality, character, and professionalism. As employees, we must interact with others in ways that puts our personality and character on display. Professionalism is
the polish with which we present our personality and character to work in order to foster a company’s goal (Shneiderman, 1990). Effective communication is an essential part of this process. Tools, such as speech recognition software, may help today’s current students acquire the skills needed to function effectively in a three-dimensional work environment.

This usability evaluation is one small, yet important step in this process of verifying the value of computing technology, specifically speech recognition software, in improving current undergraduate communications curriculum. This usability evaluation supports the software manufacturer claims that the software is easy to use and stipulates that whoever uses the product can become productive in a short period. If further evaluation and research support these claims, integration of speech recognition software into college communications curriculums should be seriously considered.
Reference List


Appendix A

Statement of Informed Consent
Statement of Informed Consent

The researcher, Professor Edward Debes, is requesting that you, the Test Subject, participate in a study of how people use computers. The purpose of this study is to provide information that will help make computers easier to use.

In this study, the researcher will first, ask you for some background information and then, second, ask you to perform some tasks on a personal computer system. The researcher has designed the tasks to be interesting and pleasant. The researcher is not evaluating you, rather the researcher is studying how easy computer systems are to use. The researcher will hold all information you provide and all data collected concerning your performance in strict confidence. The researcher will use the information for statistical and summary purposes only, and will make certain that your name is not associated with the procedures in our study.

To the best of the researcher’s knowledge, there are no physical or psychological risks associated with the procedures in our study.

The researcher will assist you and will answer any questions you have. You are completely free to stop participating in the experiment at any time.

If you are willing to participate, please sign below. If you are under eighteen years of age, a signature of a parent or legal guardian is required.

“I have read the above description of the experimental procedure and of my rights as a subject and I have agreed to participate in this study on the ease of using computers.”

________________________________________________
Test Subject

________________________________________________
Date

________________________________________________
Signature of parent if Test Subject is under 18 years of age.
Appendix B

Human Computer Interaction Usability Test Subject Questionnaire
Human Computer Interaction Usability Test Subject Questionnaire

Directions: Please provide the following information:
(If you do not have enough room to complete your answer in the space provided, you may continue your answer on the back of this page.)

First Name:
Gender:
Age:
Occupation:

1. Do you have any writing composition interests?

2. What are your primary job responsibilities?

3. What computer hardware do you have experience using?

4. What software products do you have experience using?

5. Do you have any experience using any voice recognition products?

6. Have you participated in any computer-based training in the past?
Appendix C

Software Evaluation Questionnaire
User: 1 2 3 4 5

This is a software evaluation questionnaire for Dragon Naturally Speaking Version 6.

Directions: Please fill in the leftmost circle if you strongly agree with the statement. Mark the next circle if you agree with the statement. Mark the middle circle if you are undecided. Mark the fourth circle if you disagree, and mark fifth or last circle if you strongly disagree with the statement (Kirakowski, 1994).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I would recommend this software to my colleagues.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>2. The instructions and prompts are helpful.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>3. Learning to use this software is difficult.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4. I sometimes do not know what to do next.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>5. I enjoyed my session with this software.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>6. I find the help information useful.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>7. It takes too long to learn the software commands.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>8. Working with this software is satisfying.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>9. I feel in command of this software when I use it.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>10. I think this software is inconsistent.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>11. This software is awkward to use when I want to do something that is not standard.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>12. I can perform tasks in a straightforward manner using this software.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>13. Using this software is frustrating.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>14. It is obvious that the software was designed with the user' needs in mind.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>15. At times, I have felt tense using this software.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>16. Learning how to use new functions is difficult.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>17. It is easy to make the software do exactly what you want.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>18. This software is awkward.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>19. I have to look for assistance most times when I use this software.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>20. The software has not always done what I was expecting.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
Appendix D

User One Dictation Results

**Errors = 28**

I am now able to talk to my computer!

I am now able to call to my computer!

I am now able to talk to my computer!

Dear grinning, and how was germinating. I am so glad you brought me chocolate and CUC and love Angie.

Dear Bret. How is Germany. I am so glad you bought me chocolate. We will share the treats at your party. Love, Angie

Dear Brett. How was Germany. I am so glad you brought me chocolate. We will share the treats at your party. Love, Andy

Good morning boys and girls. Today we have music and we will be reading a news story about a bearer. How many of you bought a bearer with you today? You may get Alger bearer and share it with a person next CEO. Tell your partner in your bearer’s name and how you got this special bearer.

Hood morning boys and girls. Today we will have music and we will be reading a news story about a bearer. How many of you bought a bearer with you today? You may get out your bear and share it with a person next to you. Tell your partner you bearer’s name and how you got that special bear.

Good morning boys and girls. Today we will have music and we will be reading a news story about a bearer. How many of you bought a bearer with you today? You may get out your bear and share it with a person next to you. Tell your partner you bearer’s name and how you got that special bear.
Appendix E

User Two Dictation Results

Errors = 83

I'm now able to talk to my computer.

Time now able to talk to my computer.

I am now able to talk to my computer.

dear Bret Howell's Germany I'm so glad you brought me chocolate will share the tree tenure party love and he

dear Bret Howell was Germany I'm so glad you brought me chocolate will share the tree tenure party love and he

dear Bret Howell's Germany I'm so glad you broadly chocolate will share the tree to your party love and he

Sing the morning boys and Earl's. Today will have music and we will be reading this news story about a bear. How many of you brought a bear with heat today. You make out your bearing share with the person next you. Tell your partner in your bear's name how you got this special bear

The morning boys and girls. Today we will have music and we will be beaten a news story about a bear. How many of you brought a bear with heat today. You may need to you may get out your bear and shared with the person next you. Tell your partner your bear his name how you that the special bear.

That morning boys and girls. Today will have music and we will be reading a news story about a bear. How many of you brought a bear with heat today. You may get out your bear and share it with the person next you. Tell your partner your bear his name how you that the special bear.
Appendix F

User Three Dictation Results

Errors = 47

Eye and able to talk to my computer!

Eye and able to talk to my computer!

Eye and able to talk to my computer!

Dear Bret, Howlett Germany? I am still glad you bought knee chocolate. Wheat was shared treat that you'll party. They'll, and he.

Dear Bret, Howlett Germany? I am still glad you bought knee chocolate. Wheat was shared treat that you'll party. They'll, and he.

Dear Bret, Howlett Germany? I am so glad you bought me chocolate. Wheat was shared treat that you'll party. They'll, and he.

Good morning boys and girls. Today we will have music and we will be reading a new story about a day or. How many of you bought the bandwidth to today? You may get out of LeBaron Jarratt with the person next to you. Tell you a partner your band's name how you guys get special bear.

Good morning boys and girls. To date we will have music and will be reading and the story at bat bear. How many of you bought a bear whiskey today? You may get out of LeBaron and shared with the person next to you. Tell you a partner your band name how you practice special bear.

Good morning boys and girls. To date we will have music and will be reading and the story at bat bear. How many of you bought a bear whiskey today? You may get out of LeBaron and shared with the person next to you. Tell you a partner your band name how you practice special bear.
Appendix G

User Four Dictation Results

Errors = 35

I am now able to talk to my computer!

I am now able to talk to my computer!

I'm now able to talk to my computer!

Dear Bret Hollis Germany? I'm so glad you brought mean chocolate we will share the treats at your party. Love, Angie

dear Bret, Hollis Germany? I'm so glad you brought mean chocolate. We will share the treats figure party. Love, Angie

dear Bret, Hollis Germany? I'm so glad you brought mean chocolate. Will share the treats at your party. Love, Angie

good morning boys and girls. Today will have music and you'll be reading this story about a bear. How Leyva brought a bear with you today? You may get out of your bear and share it was a person next year. Tell your partner in your bear's name and how you got this special bear.

Good morning boys and girls. Today will have music and will be reading a news story about a bear. How many of you brought a bear with you today? You may get out of your bear and share with a person next year. Tell your partner in your bear's name and how you got a special bear.

Good morning boys and girls. Today will have music and we will be reading a news story about a bear. How many of you brought a bear with you today? You may get out of your bear and share with us but with a person next to you tell your partner in your bear's name and how you got this special bear
Appendix H

User Five Dictation Results

Errors = 13

I am now able to talk to my computer!
I am now able to talk to my computer!
I am now able to talk to my computer

dear Bret, how was Germany? I am so glad you brought me chocolate. We will share the trades at your party. Love, Angie

dear Bret's how was Germany? I am so glad you brought me chocolate. We will share the trades at your party. Love, Angie

dear Bret, how was Germany? I am so glad you brought me chocolate. We will share the treats that your party. Love, Angie

good morning boys and girls. To date we will have music and we will be reading a news story about a bear. How many of you brought a bear with you today? You may get out your bear and share with a person next to you. Tell your partner your bear's name and how you got the special bear.

Good morning boys and girls. Today we will have music and we will be reading a news story about a bear. How many of you brought a bear with you today? You make it out your bear and share it with a person next to you. Tell your partner your bearers name and how you got the special bear.

Good morning boys and girls. Today we will have music and we will be reading a news story about a bear. How many of you brought a bear with you today? You may get out your bear and share it with a person next to you. Tell your partner your bearers name and how you got the special bear.
Appendix I

List of Comments Made by Test Subjects

“Why doesn’t the software talk to me?”

“Easy to follow text...not real involved.”

“Software jumps ahead and I lose my place when that happens.”

“Easy to use.”

“Learning about speech recognition while you train the software is a cool idea.”

“I would buy this product for my father-in-law.”

“Fun! Fun! Fun!”

“It takes so long for the text to appear after words are spoken.”

“Can I start over at the beginning? I misread the instructions.”

“Awesome”

“System was too slow at times.”

“The software should tell me to make the document active.”
Appendix J

Task List Distributed to Users for
Dragon Naturally Speaking

Installing and Training

Create a new user

After installation is complete and you launch the software for the first time, the New User Wizard will ask you to choose a user name. Type your name. The software automatically picks the best speech model and vocabulary* settings for your system. If you are using a USB microphone, choose the USB microphone from the “Dictation source” drop-down list. If you are using an array microphone or handheld recorder, you will also need to change the “Dictation source” setting to match your microphone or recorder.

Next, type your name. Change the “Dictation source” and “Vocabulary” settings if you need to. Then click Next.

Vocabulary

You can select from among the following vocabularies:

General: A large vocabulary providing excellent recognition accuracy for general, business, and professional dictation.

Commands Only: A limited vocabulary containing only command words and phrases. Choose this vocabulary if you want to use Dragon Naturally Speaking® only for command and control of programs and not for dictation.

Teens (US English only): A large vocabulary containing words selected for a student population and providing excellent recognition accuracy for higher-pitched voices, for example ages 11 through 18.
**Audio setup**

Now that you have chosen your user name, click Next to have the New User Wizard lead you through setting up your microphone.

**Positioning your microphone**

First, the New User Wizard explains how to position your microphone. Click Next to continue. The following sections provide detailed information on positioning the microphone. Once you have correctly positioned the microphone, click Next to continue.

**Position the microphone**

Put the microphone on your head so that the headband goes across the top of your head. Adjust the headband so that it is comfortable and feels like it will stay in place. For additional comfort, most microphones can be adjusted so that the boom extends down from either the left ear or the right ear, as you prefer.

Next, bend or adjust the boom so that the foam windscreen, which shields the microphone element from excess noise, is near your mouth. Correct positioning is crucial for recognition accuracy. Start with the microphone at the corner of your mouth, about a thumb’s width away.

If the microphone is too close to the center of your mouth, it may pick up breathing sounds, which Dragon Naturally Speaking® may mistake for “the,” “a”, or other small words. In this case, move it closer to the corner of your mouth.

The windscreen should not touch your face, hair, or beard. The proper distance from your mouth to the microphone is roughly between one-half and one inch—about the width of the thick part of your thumb.

Place your thumb between the microphone and your mouth. One side of your thumb should be touching your mouth; the other should be just brushing the windscreen. Checking this distance with your thumb is a good way to get a feel for proper microphone positioning. While you work, the microphone may shift, degrading recognition accuracy. Doing a quick “thumb check” ensures that the microphone is where it should be.

To pick up the best sound, the microphone element should be pointed toward your mouth. The microphone element, covered by the foam windscreen, is the electronic part that actually hears your voice. On most microphones a small dot just outside the windscreen indicates the direction in which the microphone hears best. Be sure that dot is pointed toward your mouth. (On some microphones, you may need to partially remove the windscreen to find the dot. Other microphones have a flat edge near the microphone element rather than a dot—point the flat edge toward your mouth.)
Consistent microphone placement is vital for accurate speech recognition. Position the microphone exactly the same way each time you use the software. If the microphone drifts, your voice will sound different to the computer and your dictation might come out garbled.

**Adjust your volume**

From the screen “Adjust Your Volume” (shown below), click the Start Volume Check button and read the text that appears. After Dragon Naturally Speaking® adjusts your volume setting, click Next to continue.

**Sound quality check**

Click the Start Quality Check button and read aloud the text shown on the screen. As you read, the computer is evaluating your sound setup. After about half a minute, Dragon Naturally Speaking® will display a “Speech-to-Noise ratio”, which indicates the clarity of your sound system. If Dragon Naturally Speaking® labels this measurement “PASSED,” you are ready to go on—click the Next button to continue. If not, see “Testing your sound system” on page 221.

**New user training**

Follow the on-screen prompts to begin the process of training Dragon Naturally Speaking® to recognize your voice.

Training takes place in two parts. For most microphone users, the first part consists of reading two sentences aloud (see illustration). Click Go to turn on the microphone and then read the sentences shown in the successive screens. The Demo button gives you an example of how to read aloud during training.

Next, the software will ask you to choose a training text to read. Select the text you prefer and click OK.

For the next few minutes, you will read passages that appear on-screen. You can take a break at any time by pressing the Pause button. Press Go to resume. You will be reading for about five minutes.

As you speak, the words you say turn from blue to black, showing that Dragon Naturally Speaking® hears them. If you stop speaking, a yellow arrow will show you where to begin speaking again. Occasionally, you will read a sentence or paragraph but the yellow arrow will point you back again to the start of what you already read. This indicates that something in the passage you read did not match what Dragon Naturally Speaking® expected to hear. Read the passage again, pausing after each sentence. If you cannot get past a particular word after saying it three times, click the Skip button to continue.

During training, dictating punctuation is not necessary. In addition, if you happen to misspeak or mispronounce a word during training, do not worry about it—just continue reading.
NOTE: When you have read enough, this message will appear:

Click OK to adapt your user files. Depending on your computer, this will take from one to five minutes or more. When adaptation is complete, follow the on-screen instructions to finish setting up your user files. The DragonBar will appear at the top of your computer screen, waiting for you to start talking. Congratulations—you are ready to dictate!

For best accuracy, be sure to complete a second session of general training as soon as possible. Do this by clicking on the Tools menu on the DragonBar, then clicking Accuracy Center. After the Accuracy Center window opens, click “Perform additional training”. This second training session will take a bit longer than the first, but the improvement in accuracy will be well worth the additional time spent.

**Training problems and solutions**

**Symptom:** Nothing changes on the screen as you speak and the yellow volume bar never turns green.
**Cause:** Dragon Naturally Speaking® is not hearing your voice.
**Solution:** Check that the microphone is attached correctly. If the volume bar still does not move as you speak, test your microphone by recording a wave file using your Windows® Sound Recorder and playing it back.

**Symptom:** The yellow volume bar changes to red as you talk.
**Cause:** The sound signal is too loud.
**Solution:** Make sure you are speaking at the same volume as you were during the Audio Setup Wizard and that your microphone is still positioned correctly.

**Symptom:** You are not able to complete the first two enrollment screens (the first two sentences that Dragon Naturally Speaking® asks you to read). This may be because you are not able to speak continuously, due to a disability or other reason.
**Solution:** Have another person read these two sentences to Dragon Naturally Speaking® so that the program can progress to the next part of training.

**Starting to Dictate**

Now that you have installed your software, created, and trained your user files, you are ready to dictate your first sentence.
First, start your word processor, such as Microsoft® Word or Corel® WordPerfect®, and begin a new document. The text insertion point should be at the start of your new document, as if you were going to begin typing.
For Dragon Naturally Speaking® to recognize your voice, you need to turn on the microphone. The microphone on/off control is located on the DragonBar, the horizontal bar that appears on-screen whenever you are using Dragon Naturally Speaking®.

Click the microphone button to turn the microphone on. The button and the volume meter next to it change to show if the microphone is “off” or “on”.

Off—the computer cannot hear you.
On—the computer is listening. It will type whatever it hears.

Read this:

I am now able to talk to my computer!

(Say “exclamation point” for “!”).

Next, take a list provided by the researcher and dictate. Remember to say punctuation. For a complete list of punctuation, see page 250.

Did it work?
If no words appear, check that the microphone is plugged in correctly and is turned on (by pressing the microphone button on the DragonBar, as described above). If only one or two words appear, your microphone may be set up incorrectly or your sound settings may need to be changed. See page 221 for instructions on how to test your sound system.
If words did appear, but many are not what you said, that is okay. Dragon Naturally Speaking® will improve as you use it.

Note: For a new paragraph, say “New Paragraph” run together as one word, with no pause after “new.” For a new line, say “New Line.”
After dictating a paragraph or two, go back and correct any recognition errors with the keyboard and mouse, just as you would usually correct errors in a word-processed document. You will learn to correct by voice in chapter 4, “Correcting and Editing”.

This ends the research exercise. Thank you for your participation

*About vocabularies
A vocabulary in Dragon Naturally Speaking® is a body of information that includes a word list and a language model. The word list includes information about all the words that the program can recognize. The language model contains usage information about those words. Dragon Naturally Speaking® uses a vocabulary to recognize words correctly based not only on the sound of the words, but also on their context. When you create a new user, you select the vocabulary on which to base the user. When you create a new user or new vocabulary for an existing user, you select the vocabulary on which to base the vocabulary you create.