Fall 2008

Voice Recognition Software: A Brief Case Study

Kirstin Duffin, *Eastern Illinois University*

Voice Recognition Software: A Brief Case Study

By Kirstin Duffin, University of Wisconsin-Madison

Dragon NaturallySpeaking is voice-activated software that allows users to communicate with a computer by means of their voice rather than their keyboard. The software's aim will make anyone who failed Keyboarding 101 giddy with delight. Forgo typing, be it for a document, an e-mail or instant messaging. Dragon listens to your voice and records your composition faster than you can type. After a brief, 15-minute training and tutorial session, during which Dragon listens to the user and he or she learns commands to navigate within Dragon, you are ready to start using the program.

Applying the magic of Dragon NaturallySpeaking to the University of Wisconsin-Madison's Oral History Project, the staff wanted to determine whether a student user of Dragon, reciting a narrator's words to the computer, would be a more cost-effective means of transcribing audio interviews as compared to typing a transcript.

Starting out, the brevity of the training surprised me. After hearing myself read just a few paragraphs, I wondered how accurate Dragon would be. The user can opt for further voice recognition training with Dragon, but as my

time with testing the software was short, I bypassed this choice. I wanted to get started, commanding Dragon with such powerful keywords as "scratch that" (deletes what was just written), "new line" (acts like the return key, creating a new line), and "go to sleep" (hibernates the program—in this mode, Dragon will not register audio through the microphone). After completing the training, I felt prepared to use the program.

Despite the succinct tutorial, Dragon amazed me with its accuracy. It understood words even when I slurred my

Emerging crisis grant to aid study of displaced Afro-Colombians

By Charles Hardy III, OHA President Elect

The Oral History History Association has awarded its 2008 Emerging Crisis Research Award to Ann Farnsworth-Alvear and Carlos Rosero of the University of Pennsylvania.

In 2006 the OHA established an Emerging Crisis Oral History Research Fund, which awards an annual grant of up to $3,000 to enable researchers to undertake oral history research in situations of crisis in the United States and abroad. This year, the fund attracted more than a dozen proposals, many of excellent quality. Indeed, so many of the proposals were clearly fundable that the committee found the selection of the winner problematic.

After serious soul searching, the selection committee, composed of Patrick Carlton at the University of Las Vegas, Jeffrey Gerson from the University of Massachusetts at Lowell, and Debra Schultz of the International Center for Transitional Justice, gave the award to Ann Farnsworth-Alvear and Carlos Rosero of the University of Pennsylvania for a project entitled "Twice Refugees: Afro-Colombians and Political Action During the Explosion of Violence Along Colombia's Pacific Coast."

Farnsworth-Alvear and Rosero will use the award to travel to hamlets along the Yurumangui and Guapí rivers where they plan to videotape oral history interviews to improve understanding of the emerging humanitarian crisis facing black communities along Colombia's Pacific Coast. By early 2008, Colombia had more than 4 million internally displaced persons (*los desplazados*), including 1.8 million Afro-Colombians, the majority of whom come from the Pacific Coast. In this region they have been caught in struggles between the Colombian military forces, regionally-based guerilla armies and local paramili-

tary units. Drug trafficking fuels the crisis, with the armed actors all having links to the Pacific Coast import-export trade in coca and small arms. These displaced persons have migrated to Buenaventura, Cali, Bogota and other cities where they have struggled to survive and to bring international visibility to the refugee crisis.

To provide depth to this humanitarian crisis, and move it beyond the focus of victimization, the investigators plan to conduct oral interviews with *desplazados* and with activists who have remained in the rural communities. These interviews will elicit their memories of political struggle before and during the recent wave of violence. The researchers also will interview persons who have been involved in attempts to establish peaceful communication and relationships with guerilla and paramilitary leaders while avoiding confrontation with the national government's armed forces.

The investigators will work with the Latin American and Latino Studies Program and the Van Pelt Library at UPenn to create a digital archive housed at the university. They also plan to publish some of the interviews in English and to place translated materials into the hands of African-American journalists interested in gaining an understanding of the emerging crisis faced by Afro-Colombians.

The Emerging Crisis Research Fund has quickly become one of the OHA's most important initiatives. Due to the high quality of submissions, the OHA will be looking for funds to increase the number of worthy projects it can assist each year. Those interested in contributing to the Emerging Crisis Fund or willing to assist in fundraising should contact the OHA. ❖

Voice Recognition Software: A Brief Case Study

speech, sacrificing my enunciation while trying to understand the narrator's words. Dragon knew to put words in context. For example, if I said, "They're commuting over there using their bicycle built for two," Dragon would understand when to spell they're as opposed to there or their. Also, although the words we're and were can sound alike if one were speaking quickly, Dragon could distinguish between the two.

Transcribing using one's voice rather than one's fingers is a new skill that takes time to master. If one can accept and expect a steeper learning curve, I believe Dragon can be used effectively.

The program had difficulty with proper nouns. Sometimes it wrote a more common word with a similar sound, and sometimes it guessed admirably at the spelling of the name. Command words caused confusion for Dragon within the context of the interview. Saying "edit" could bring up the edit menu, pausing the transcription; dictating "period" could instead type the punctuation mark. These ambiguities aside, Dragon's precision was impressive.

On the other hand, Dragon's "autopunctuation" feature worked less satisfactory. With autopunctuation turned on, Dragon will insert commas and periods with the natural pauses and inflections of the speaker's voice. Less common punctuation marks, such as the colon, question mark and exclamation point, must be inserted by the speaker. With autopunctuation off, the speaker must instruct Dragon where to place all punctuation, so the transcriptionist would need to remember to say "period" and "comma" in the correct places. Otherwise the transcript would be a string of text with no punctuation whatsoever.

I tested the software with autopunctuation on and learned that transcribing does break the normal flow of speech. While I dictated the oral history narra-tor's words, I focused on what was being said in the interview rather than on my own intonations, which caused Dragon to add commas and periods in strange places. Transcribing with Dragon required me to focus on audio output rather than the interview's context, and the product was a transcript with punctuation placed sporadically.

There are some practical ways, however, that I could make Dragon work to my advantage. Turning the headphone feedback of the transcriptionist's voice off, for example, stopped my voice from covering the narrator's. The "add new command" feature was a hidden treasure. This allowed for the creation of self-formatted shortcuts.

Two such shortcuts were especially helpful. When I came to an unclear spot in the interview, I created the shortcut "sounds unclear" (I got to choose the catchphrase). Saying it would cause unclear to appear. Instead of having to say, "unclear, select unclear, bold that," this shortcut let me bypass vague dialogue with ease. To identify the start of a new speaker, I was having trouble getting Dragon to understand that I wanted it to type letters. Saying "MS," for instance, might yield "aim pass," among other combinations. I trained Dragon to understand that when I said, "MS," it should output "MS:" Without this shortcut I would have to say, "cap m, cap s, colon," an inefficient and frustrating technique.

After the initial dictation, I tested Dragon's proficiency at helping me edit the transcript. As I had limited time to work with Dragon, using it to help clean up the text did not work well. With Dragon, for example, if the sentence ended with the word orangutan, I needed to say, "Insert after orangutan, period." With a number of commas and periods to add and delete, the use of a mouse appeared to be a more efficient method.

After working with Dragon for a few hours and becoming more familiar with its operation, I grew attached to it. Using Dragon, it seemed, took more practice to overcome the learning curve as compared to typing a transcription. Once I established a rhythm, I could not imagine typing faster. I must add two caveats. First, the transcriber can work only as quickly as he or she understand the narrator's words. After hearing the interview, the transcriptionist must generate a reiteration either through vocal output (if using Dragon) or finger movement (if typing). We communicate using the keyboard nearly every day of this modern age; at typing we are highly practiced students. Voice-activated writing is a new skill with which many, including myself, have no prior experience. If I improved as much as I did during my 20 hours with Dragon, I can imagine someone with daily practice could do well with it.

It may be easier to type over the voices in the interview, perhaps because the brain can more readily process the conversation while typing. After training and working with Dragon, however, I believe using it would save time over typing. It would take practice to learn to talk over the interview's voices, to be sure, but one can acquire that skill.

In my brief experience with Dragon, as long as I could understand the voices in the interview, I could shadow the conversation. I could repeat what was just said, following the interview in this manner for considerable stretches without pausing. I could not type faster than Dragon dictated. I did struggle, however, to understand the narrator or interviewer, detaining me at times from making steady progress. Not being able to hear the words would cause anyone, whether typing or using the Dragon software, excessive toil. With time anyone could grow quite adept at decoding mumbled utterances and understanding voices through poor audio quality.

Transcribing using one's voice rather than one's fingers is a new skill that takes time to master. If one can accept and expect a steeper learning curve, I believe Dragon can be used effectively. There may be some interplay between the use of Dragon and the keyboard when it comes to editing the transcription. Overall, student transcription using Dragon NaturallySpeaking seems a plausible way to transcribe interviews for the university's oral history program. ❖