

Hands Free

Steve Hughes 195 North Mill court, Atlanta Georgia 30328

This short newsletter is being done using the new IBM ViaVoice force recognition software. I have spent considerable time setting the software up exactly according to manufacturer's specifications. I have trained, using a full two hundred and sixty-five word script, exactly as recommended by itself or package. I am dictating into IBM's SpeakPad software application. In short this should be about the best possible test of this software is ability to recognize my voice. I must say I am disappointed. The recognition accuracy of the software, a Mason taken by itself, falls far short of the advertising. IBM's claims for the software would have you believe it is adequate to use in normal office environment. This simply not true. Even the smallest amount that Crown or is tremendously be grades recognition accuracy.

That last sentence should have read, this simply not true, even the smallest amount of background lanes tremendously grades the recognition accuracy. I am not intending to make any corrections to the spelling this document before printing. In some cases it's going to be fairly difficult for the reader to figure out what I was trying to say. Obviously this off word is not known word slide, "sspk" and cannot be expected to know word like zine. I have tried to train some of the common words used in the end fan the right thing into the system with lukewarm success.

As I look over the 2

So I have already dictated, I'm sure everyone is foreign to or contract to the original scenes that I used to produce cut directly on been no formal stencil with their incredible high number of titles. In that sense I suppose I cannot criticize the software to heavily for making the same kinds of mistakes that I would once of made typing-Material. However I'm never considered that can be satisfactory. I notice occasionally the package will throw when a new
Or even a

Without of my requesting. It also sometimes drops words or phrases. For example descends before the last one should have read "Without my requesting yet". OK that came a little bit closer to what I was trying to say. I believe this package could become usable, the f you will willing to learn how to speak to react. The exact way in which words are pronounced and how they're pronounced in combination with other words has a profound effect on the program's ability to recognize those words. I had considered simply reading a few paragraphs from short story, or something of that sort, but decided that a more normal free flow dictation of the type that one might use in a newsletter would be more appropriate.

I have not decided if I will typed a corrected version of the same text so that you will have something to compare this version with. In one sense it would give you a better idea of the mistakes the programbut it would be rather tire said. Sometimes the program

complaint confuses two or three words and treats them as a single word. It will also from when extraneous words just out of the blue.

The factors that seem to affect recognition strongest or: microphone placements, background noise and exactly how each word is pronounced. Background noise can be hard to control, for example when the air conditioner turn sell the programs and has a problem. A plane flying overhead can generate enough new ways to affect recognition. Microphone placement is sold delicate that I have to do if you test sentences every time I moved my head or take a drink of water. As I sat here dictating to this program I'm learning to make it recognized some words that it originally had trouble recognizing.

This later problem can be helped by going yen and identify and to the program words did has gotten thrown an corrected. The program will save those corrections and in the future to a better job of recognizing the words that were collected. It maintains a database for each speaker the news is the program and the corrections that they have made.

As critical as I am being of the software, I have to the minute I'm absolutely amazed that it works as well as it does. Continuous speech recognition, recognizing word without pauses for breaks between them, is one of a heart is software problems we have tried to tackle in the computer industry. Boris recognition has been commonplace in movies and TV shows for many years. So, and that many people assume it is easy to do. In fact it is an incredibly difficult programming/understanding problem. The main difficulty come from the fact that we humans do not understand fully out we hear. There are many theories to explain how our own brains beat recognize speech. Also we humans do not do a very good job of recognizing the MedSpeak.

I once did the test in my old job by having a person mass grave as a machine unrecognized in human speech most people flocked that she did not do a good enough job to use in the real world. But when

You will notice that the microphone shifted on my head slightly and last paragraph of last few sentences of the previous paragraph have a much higher number of years in. Now I have read positioned microphone and it is performing better.

Back to my story, I set up a very good human telephone operator to listen to people entering speech to what they fault was a computer program. She then responded to their commands through a keyboard. Everyone contested the program told me it was not good enough to be used in a real application. The person doing the test for me was one of the best telephone operators into our VoiceCenter. This highlights the fact that we expect much more from machines than we do people. If we see that a person has not understood what we said we will repeat yourself or say the same thing in different words. We expect the machine to never make a mistake.

This system is definitely the best was recognition of Henderson. Still high load extremely unhappy using it would be heard from for any critical application. All I have just deposition the microphone you again since I solve that the accuracy was going with a

damned. All my son of the sentences are simply not the winning anything to person reading them. The report depicted still I think gambling put this through untouched so that everybody can see exactly how bad is the bid for exactly how good it can be spending all your viewpoint.

If you want to play with toys recognition I would recommend IBM's ViaVoice product. For its costs which is only about 705 dollars, it is really a very nice,. That last dollar amount should have been an 705 not seven hundred and five dollars. Let's try it one more time, 705 dollars. I can't seem to get this program to recognize the fact that this but the dollar amount of trying to say is not seven hundred and five dollars but only 75 dollars. There are finally got it to recognize the dollar amount. -Pope with

I'm also finding that ViaVoice tires rapidly speaking in this even measured tone so I think that this zine is about done and you guys will make of it what you can. Perhaps in the future, as I choir of other was recognition packages, I will do the same kind of newsletter using the new packages. If I make any real progress in getting this package to work all we do this exact same zine.

I hope you can make enough sense of this text to lead least understand why was turned aside. In many cases it has produced totally unrecognized of all sentences still I was tried tested and that half.

I'm going to print and mail this immediately before I start making changes to try and make it alone more readable. There is a strong temptation to cleanup the places where the program put on wanted new lines and unwanted paragraphs as well as extra words. But doing that would not be fair to the test inspiring to do. So low what you're reading is exactly what the program produced. And at least I have something in the next mailing.

See you guys at the convention.

A Few Comments on the Test (input the old fashion way)

I'm a big fan of voice recognition systems. Probably because I've always wanted to say things like "house – lights on" and have the house pay attention. Over the years, I've worked with people like Bell Labs testing their systems and the technology has always been "5 years" away from being really practical. For many years the excuse was computer power, the Bell Labs guys used to say they needed the equivalent of a main frame computer to make their software work. Well, the little test I just finished was run on a 266mhz Pentium II which has several times the processing power of one of the old mainframes. The problem with voice recognition is that speech recognition is not the problem. Voice recognition requires *understanding* speech, not just recognizing sounds. A large part of the time when we think we're hearing the words someone says we're actually hearing the *meaning*.

When I showed the test pages to Suzanne, she proceeded to read one of the garbled paragraphs to me. It made perfect sense! Even though the words were wrong the *meaning* was intact. Don't believe me? Try having someone read a couple of paragraphs of the test out loud to you in a normal tone of voice. Suddenly the "wrong words" will start to make good sense. We convey meaning using speech in many ways: syntax, inflection, logic, context and previous experience. The problem with computer voice recognition is that most of that is missing. I say most because modern programs understand syntax and even a little logic. In time, say *five years*, we may get usable voice recognition software. But I doubt it, I think we're going to have to make some major advances in understanding human thought before we crack this problem.

Does that mean the house will never understand my commands? Of course not, that's a much simpler problem, I only have to teach the house network to understand a few discrete commands in my voice to implement voice control. In fact I'm working on it now.

One last thought, if anyone had told me 10 years ago that I could sit in my home office and dictate to my computer using a \$75 program and get anything even half as good as the test text, I would not have believed them. *ViaVoice* and its competition from Dragon Systems are marvelous pieces of software. Part of the problem I'm having is probably my own accent. If any of you have tried this package, or others like it, I would like to hear what you think about them.