

## Custom Speech Recognition and Language Processing Developers for Applications and Software



**Jeff Adams**  
Founder & CEO

**CEOCFO:** *Mr. Adams, what is the idea behind Cobalt Speech & Language?*

**Mr. Adams:** Lynn; that is a good question. We do custom speech and language processing for a variety of customers in a variety of industry segments. The idea is that many companies out there want to do something involving speech or language or both and they find that it is a challenge. It is not an area where you can just hire your average software developer and say, "Here build us a speech recognizer that will do these five things." It takes real expertise and specialized knowledge. Therefore, right now I have a team of about ten who do have experience in the field. Customers hire us and we build whatever it is they need. It might be an application or a service or some back-end analysis tool or something like that.

**CEOCFO:** *Are there any sort of off the shelf solutions or is the entire arena customized?*

**Mr. Adams:** There are off the shelf solutions. You can go to a company like Nuance or Interactions or Sensory; there are a handful of companies out there that will provide speech and language technology. However, generally they offer sort of generic speech recognition or speech synthesis. If all I want is to recognize the speech of someone describing their day at the beach, for example. I could probably get something from Nuance that would do that reasonably accurately. Many of our customers are looking for something more specialized that is not served by off the shelf offerings.

For example, one of our early customers had developed their own specialized laser optical microphone that could detect vibrations in a person's cheek and neck. They wanted us to help them build the speech recognition technology that would use that microphone. They did not have the option of going to one of these other providers and taking something off the shelf. Those providers generally take existing technology and they make it available in a fairly generic way, whereas we will go in and say, "Ah! We will reinvent the whole process to handle your particular needs." I will give you a couple of examples. We have another customer who is working with pharmaceutical companies in evaluating drug trials. As part of their process they have medical professionals that will interview the subjects in the drug trials. They will interview them on a weekly basis and see how they are doing. This is especially useful for mental health medications for treating depression and schizophrenia. A doctor might interview the patient once a week. Now, these doctors are not in a traditional doctor patient relationship and they have different motivations and there are different pitfalls that could come up. These doctors are just getting paid by the insurance company to do an interview to evaluate the medicine. As a result, sometimes they are not as diligent as they might be if they were the patient's real doctor. They might skip questions or not follow up where they should and so on. Therefore, this company has engaged us to automate the process of evaluating those interviews. Those interviews are recorded and then we, using speech recognition and analysis of what happens; we can go in and see if the doctor is engaged, if they are following up where they should and it requires a level of analysis that frankly you just cannot get from an off the shelf product. Therefore, we have to go in and design it from the ground up.

**CEOCFO:** *What are some of the challenges in doing this?*

**Mr. Adams:** At a very high level, I will not dive down into the weeds of the technology unless you twist my arm and ask me to, but at a very high level we have a sort of evangelical mission to democratize speech technology. Many companies have just been shut out from access to the kind of PhD computer scientists that have expertise in this field. They just do not have access to them, because those guys are all employed by either universities or governments or big companies like Microsoft, Google, Amazon, Facebook, Apple and so on. Your smaller company just does not have access to the technology. Therefore, we have this evangelical mission of bringing together expertise on the one hand with the industry. When I say the industry I mean all industries--anyone who wants access to this kind of technology. In a way, we are a

middleman. We bring the experts together with companies that need that expertise. There are challenges on both ends of that. On the one hand it is challenging for us to find companies that need the technology, because it is esoteric enough that they do not always understand what they need. They do not understand that they cannot just get it off of the shelf. They do not understand what the potential is of what they could do with it. On that side our challenge is finding customers, helping them understand what the opportunities are and then persuading them to work with us to put something together. Then on the other hand it is really challenging for anyone to hire speech and language experts these days. They are in high demand. The demand is much greater than the supply, so there are bidding wars in salaries. There are the same six or seven companies that are trying to hire the same folks. It is really hard for us to find the experts and persuade them to join us. We have our ways of doing that, so we have a pretty respectable team of, as I said, ten folks right now, and we are growing. We have a few more scheduled to start by the end of the year.

**CEOCFO: *What have you learned in the year since you have founded Cobalt Speech & Language? What is different in what you do today?***

**Mr. Adams:** Yes, and it has been a year! Tomorrow is our birthday! This is my first foray into business. I am not a serial entrepreneur. I have never done this before. I have managed speech teams at a bunch of different companies; most recently Amazon where I led the group that built the new Amazon Echo. Previously I had worked with the teams that built Dragon NaturallySpeaking and some other products that people may have heard of. However, this is the first time that I have been involved on the business side. There is not really another company like this that we can follow. We learned a lot just in terms of how to come up with a good business model and how to approach the market and how to organize ourselves to be able to support multiple customers at the same time, and how to organize our own processes internally to make sure that we are giving all of our customers more than what they asked for and delivering on our promises. The main thing that I would say that I have learned is just how to go about this new kind of business. We are sort of inventing a new way of working, at least in this industry.

**“We have a sort of evangelical mission to democratize speech technology. Many companies have just been shut out from access... We bring the experts together with companies that need that expertise.” - Jeff Adams**

**CEOCFO: *Is there anything that you cannot make a program for? Are there certain things that just cannot be done or is there a solution, it is just a matter of having the brain power to find it?***

**Mr. Adams:** Yes. Actually, this question that you ask is key. This is one of the difficult things for us to talk to all of our customers about. People get the impression and they say, “Siri works on my phone,” or “when I call and talk to my bank or airline, I have a computer that understands me.” People get the impression that speech technology is a solved thing; that it just works and it is just a matter of plugging in the right thing somewhere and it will just magically work. I am going to focus on speech recognitions in my answer here, but the same thing applies to language analysis or speech synthesis or whatever.

If we focus on speech recognition you can look at a handful of different ways in which it can be hard or easy. I will give you five examples of that. First, on the easy side you could have clear speaker without an accent or on the hard side you could have someone who does not speak so clearly; perhaps English is not their native language or they are young or they are very soft spoken. Second, you could have a very clean, quiet environment or you could have a very noisy environment. Third, when a person is speaking we might have the microphone close to their mouth so that we get a good, direct recording or it could be across the room, in which case we get echoes bouncing off of all the walls and windows before we get it to the microphone. Fourth is the complexity of the language that we are expecting them to say. On the easy side you could have a very predictable sentence like “yes or no”, or a number. On the hard side, you could be expecting any kind of speech, such as a person reciting their life story. Finally, you could have just one person speaking or you could have a bunch of people speaking at the same time.

In these categories, right now speech recognition is good and useful as long as most of those categories are on the easy side for the particular problem. We can handle one or two of those being hard, but we could not handle all five of those things being hard. Therefore, in the case where you call up and you are talking to the airline it is one person speaking. Your mouth is close to the phone. What you are going to say is very predictable and you are probably not speaking with a strong accent. We can handle it if you are speaking with a little bit of an accent, because that is the only thing that would be hard in this case. On the other hand, the thing that we cannot do well right now, that is considered by many to be a sort of unreachable Holy Grail right now, is if you and I and three of our friends were sitting around a table and we put a phone in the middle of the table recording our conversation and then afterwards we wanted to make a transcript of that discussion, that is just out of reach for any practical, useful kind of application. That is because it is wide open what

people are going to talk about, it is a room where people are not close to the microphone and it is very spontaneous, with unpredictable speech, there are many speakers and that is really hard to do. It is hard to get more than one out of four or five words correct in a scenario like that.

**CEOCFO: *Are you funded for the steps you need to take to gain recognition? Is that an issue for you?***

**Mr. Adams:** It is not really an issue. We decided early on not to go after any funding, so we are entirely self funded. That means that we are funded by our customer's projects. We take on projects from our customers and then we hire the folks to work on those projects. As a result, we are completely self-owned. I am the sole owner of the company and we have more complete control over what we are trying to do. Early on I wanted to talk to investors about VC funding, or any kind of external funding and I quickly realized that that always came with strings; that someone said, "Okay, I will fund you, but I want you to focus on talking toasters," or some particular specific application where they thought the market was going. We did not want to give up control of our destiny in that way. We wanted to be able to decide ourselves whom to work with and which job to take on and so forth. Key to us is that we are not a product company. We do not produce products. We produce the technology and supply that to our customers and they produce products.

**CEOCFO: *What is ahead? How will the next year be different?***

**Mr. Adams:** Just in the last month or two we have had an explosion of interest from a variety of customers.

**CEOCFO: *Could you tell us why?***

**Mr. Adams:** I am not sure why. Three months ago I hired a VP of business development. He has been working hard on getting us some visibility. We went and participated in a trade show. There is a trade show every year called SpeechTEK and most of the companies there were offering "one size fits all", out of the box, plug and play kind of speech applications. However, many of the folks that were looking for speech technology needed something custom and we were the only ones there that could help them. Therefore, we got a lot of interest from that show. Much of our interest is also coming from referrals. Sometimes it just takes time for people to get used to the idea of outsourcing key technology or working with us and it is by word of mouth that that is spread. Therefore, business has been growing slowly but exponentially over recent months, the last two or three months in particular. We were reviewing our list of potential customers and this list had twenty potential customers on it six months ago and it now has sixty and they seem more serious and more aggressive about moving forwards. Therefore, I think that in terms of what the next year holds, I expect that we will grow from ten to twenty or thirty clients. I expect we will grow in the number of projects and I expect that we will start to have our first repeat business. We have a number of customers now that are an introductory kind of project. Over the next twelve months I think we will start to see some of those converting to the next generation. We will be saying; "Let us see if we can build on that."

Interview conducted by: Lynn Fosse, Senior Editor, CEOCFO Magazine

---



**COBALT**

**Cobalt Speech & Language, Inc.**  
**[www.cobaltspeech.com](http://www.cobaltspeech.com)**

**Jeff Adams**  
**978-399-9124**  
**[jeff@cobaltspeech.com](mailto:jeff@cobaltspeech.com)**