Posted on Mon, Mar. 21, 2005

Voice recognition: They speak thereby they brand
San Jose Mercury News

By Nicole C. Wong
Mercury News

Joann Landers, a Yellow Cab driver from San Jose, calls Cal North four or five times a week as she zips passengers to the San Francisco International Airport. Should she take Highway 101 or Interstate 280? Her friend always knows the quickest route.

Sometimes, she's shocked when he warns her about traffic jams and lane-blocking debris. ``Are you kidding?'' she scoffs. But his calm and soothing voice helps her keep cool.

North has no idea who Landers is.

That's because he's a voice-activated recording created for the Metropolitan Transportation Commission by Nuance Communications, a Menlo Park speech recognition company that's leading the pack in this type of technology. Nuance envisioned a retired Highway Patrol officer whose deep, warm voice never sounds rushed or flustered as he informs time-taxed travelers about 10-car pileups and clogged roadways.

As more companies and government agencies turn toward interactive speech software to manage call-center costs and satisfy clients, they're putting increasing importance on imaginary phone operators with pleasing personalities. These customized characters exemplify how technological innovation has opened new avenues for branding -- conveying the essence of an organization.

Branding has evolved with the printing press, radio, television, Internet and computer applications.

``With technology, brands have to express themselves in far more complex terms than they did even 10 years ago,'' said Peter Sealey, co-director of the Center for Marketing and Technology at UC-Berkeley's Haas School of Business. ``I'm spending time with the brand whether I'm buying it, consuming it, seeing it or listening to it on the telephone.''

The basic technology of speech recognition search engines has evolved rapidly in the past few years; their ability to understand utterances has improved from being barely adequate to often accurate.

``There was a pretty steep hill to climb in convincing people that it worked,'' said Moira Dorsey, senior analyst of customer experience at Forrester Research.

When Nuance started 10 years ago, its speech engine ran on a refrigerator-size machine and could recognize only 100,000 utterances. Now, powered by desktop PCs with better computing capabilities at lower costs, its engines can swiftly match the sound waves of whatever a caller says to the electronic patterns of 25 million words and phrases stored in its database.

The engines also are succeeding more in understanding what callers are saying because of new ways to eliminate background noise and automatically update word-recognition databases with synonyms.

With the technology much improved, the company also can focus on the next step: the personas, the shiny chrome covering the search engines.

Until a year ago, companies switched from live phone operators or touch-tone menu trees to automated agents mostly to slash expenses -- generally saving about $6 a call, according to Forrester Research. Those companies usually pulled a stock persona off Nuance's shelf, caring only whether they chose a male or female voice.

The right persona
• Firms breathe life into user interfaces

But now the process of picking the right persona has become a business in itself at TuVox in Los Altos, BeVocal in Mountain View and Tellme Networks in Mountain View -- companies that breathe life into user interfaces built on top of Nuance's speech engine technology. That's because the wrong persona is worse than no persona, experts say.

``You want to make sure you're not going to do something that breaks the brand,'' said Rebecca Nowlin Green, a Nuance user-interface architect and voice coach. ``If the brand is conservative, you don't want to throw a chatty, teeny-bopper voice at it.''

At Nuance, figuring out a voice that fits begins with two weeks of interviews with the client company. What marketing messages are communicated through the company's other channels? Is this phone system supposed to sell additional products or shorten the time live agents handle calls? How old, educated and tech-savvy is the typical

caller? Are calls usually placed from home or from a car? What's the average customer's mindset?

``Sometimes they might be calling an application to get information, in which case the persona might be more snappy and efficient,'' Green said. ``In another case, they might be calling to do a transaction, in which case it might be more important to have a persona that's patient.''

Often, companies can't articulate what they want.

So, Green said, ``if the customer has an ad agency, we'd want to talk to them because they're the keeper of the adjectives.''

The persona producer also visits the client's Web site to scout out its word choice, colors and style.

``You wouldn't necessarily throw a really bright, sunny persona at a company whose Web site is brown,'' Green explained. ``At the same time, if you have a Web site that's really snazzy, you wouldn't want to give it a drab voice.''

Crafting the right persona also requires understanding callers' socioeconomic status and ``social dialect'' so the automated phone attendant doesn't put them off with hoity-toity terminology.

All that research produces a persona that is introduced to the client as a voice on a demo tape packaged with a phony résumé, fake job references and a pretend personal statement on why the candidate wants to answer the phone hot line. The concocted personal information helps clients understand what they're buying and prepares voice actors for their roles.

Nuance invented Jenni McDermott to read e-mail messages over the phone for Yahoo's customers, using technology that converts text to speech. The design team envisioned her as a 5-foot-5, 108-pound, brown-haired, blue-eyed barista (a professional espresso maker) who had a hard time finding a job after graduating from the University of California-Berkeley with a bachelor's degree in art history. McDermott, born Aug. 20, 1979, spends her free time painting and scuba diving. She likes independent films and dislikes smokers.

The 739-word biography also details her father's job as an electric company lineman, her weekend trips to Big Sur with her dog, Brindle, and the new love of her life -- a jazz musician named Rob.

The description translated into an expressive voice that's supposed to sound inviting, yet not chatty.

The affable Gracie
• Positively chummy, has a spiky haircut

Kathryn Nymoen, a voice talent from San Jose, roams even further into the fantasy realm when she lends her voice to various personas: ``I imagine all the way down to what they drive.''

Greyhound hired Nymoen to play the part of the affable and earnest Gracie, (who, like Barbie, has no last name). The petite, positively chummy 20-something Southern gal sports a spiky pixie haircut and bats her big green eyes.

Washington Mutual gave Nymoen the role of Laurie Woods, a New Yorker in her early 40s with brown hair and blue eyes. She's efficient, educated and conservative.

That's where the persona producers stopped, and Nymoen stepped in.

``Gracie drives a Mini Cooper, and Laurie Woods drives a Chrysler Concorde,'' said the fortysomething New York native, who chauffeurs her three kids in a Honda Pilot.

Those details help Nymoen remember to constrict her throat, which raises the pitch of her voice, when she records the times and locations of departing Greyhound buses, and to utter everything with so much enthusiasm that the words mush together. When she reads the streets and cities where Washington Mutual's branches are located, Nymoen speaks with can-do professionalism at a metered pace.

Nymoen earns $500 to $2,250 an hour to do the work of recording phone prompts. Sometimes she's stuck in a small, dark, unventilated sound booth for a three-hour session reading a script filled with more than 1,000 things to say. The worst, she said, is when she's rattling off a boring list of stocks and funds. Or when she gets so sweaty that the black plastic headphones melt, leaving marks on her ears.

It's also hard striking the right emotional key when you're only hearing half of a conversation.

A few weeks ago Nymoen recorded new prompts for Wells Fargo. ``I'm sorry, status information for year-end tax information is temporarily unavailable on this system,'' she said into the microphone.

Green, the voice coach, suggested a second try: ``Can you sound more sorry? Thank you.''
Contact Nicole C. Wong at nwong@mercurynews.com or (408) 920-5730.