FROM
SPEECH STRATEGY NEWS
NOVEMBER 2012

Interview with Amy Neustein, Linguistic Technology Systems

The 'In Touch' Personal-Assistants: Next-Generation Emotionally-Intelligent Mobile Devices

Amy Neustein, Ph.D., CEO and Founder of Linguistic Technology Systems, was interviewed by Bill Meisel in late October. Dr. Neustein is Editor-in Chief of *International Journal of Speech Technology*, Series Editor of the *Springer Briefs in Speech Technology*, co-editor, with Judith A. Markowitz, Ph.D., of *Mobile Speech and Advanced Natural Language Solutions* (Springer, 2012), editor of *Advances in Speech Recognition: Mobile Environments, Call Centers and Clinics* (Springer, 2010) and co-editor, with Hemant A. Patil, Ph.D., of *Forensic Speaker Recognition: Law Enforcement and Counter-Terrorism* (Springer, 2011).

Can you tell us about the germination of your idea for an emotionally intelligent personal assistant?

I was puzzled by the fact that as smartphones increasingly become "smarter" at understanding natural language input, there doesn't seem to be a drive to make these devices "emotionally" intelligent.

But what about Siri? Doesn't Apple's personal assistant endear itself to the user?

"Cute" doesn't necessarily mean emotionally smart. In fact, such devices that engage users in carefree conversations may appear, albeit speciously, to respond on an emotional level. However, when the smart device produces a preposterous response either because of an out-of-vocabulary word that was not understood by the personal-assistant model or the task requested of the mobile device was unfamiliar to the system altogether, the bubble begins to burst.

Please tell us what you would do differently.

I am introducing the *"in touch"* personal assistant that will sense the mood of the user and perhaps in extreme cases even advise an agitated user not to get behind the wheel of a car. In contrast, if the personal assistant senses the user is in a relaxed mood, it may try to engage them in a bit of cybershopping.

Are there other tasks that can be performed by an emotionally intelligent personal assistant?

Yes. The emotionally intelligent personal assistant can be useful for e-commerce. Because so much of what is said online about products and services can be hidden in the descriptive portion of online consumer reviews, the emotionally intelligent smartphone would locate and extract the hidden attributes, both positive and negative, contained in opinion-related postings. In practical terms what this means is that the savvy smartphone would dig deep into the morass of user-generated content. In so doing, it would extract from those long-winded, and often circumlocutory, consumer postings appearing on such sites such as Trip Advisor, Overstocks.com, or Zappos.com, the more emotionally subtle features of such reviews.

So you believe your social media mining program would find the more nuanced features buried in consumer reviews?

Yes, that's it. And by bringing these nuanced features to the attention of the smartphone user, they would be given more of an insight into the products and services they are exploring on the web. Similarly, enterprises may benefit from emotionally-intelligent mining of online postings to learn how to improve their products and services.

Will this be implemented as a wholly integrated system or as an Application Programming Interface for an existing smartphone?

I've thought about that quite a bit. At this point it's best that I leave those options open. It could go either way. However, I have selected a name for this emotionally sensitive personal assistant: "AMY" (Affective-Mobility Yardstick). Using a novel natural language method known as Sequence Package Analysis (or SPA),

AMY will gauge the emotions of the smartphone user as well as mine opinion-related reviews for emotional data.

Please elaborate on your work on SPA, your natural language method intended to make the personal assistant emotionally smart?

Certainly. For the past decade, my research on SPA has appeared in peer-reviewed journals and in refereed conference proceedings, catching the eye of a number of AI researchers. For example, those interested in data mining in call centers focus on SPA's potential to "caption the text"—that is, to find subtle features in call-center recordings such as "early warning signs of customer frustration." Other AI researchers have pointed out the utility of SPA for applications other than call center operations. Those researchers have envisaged SPA as part of the broad spectra of "medical natural-language mining tools" that may assist in the successful classification of "affective" versus "informative" content found in health-related web postings.

So your work has focused on identifying and measuring the affective content of natural language data?

Yes, both spoken and text-based natural language input. Because natural language—both spoken and text-based—is inextricably tied to human emotions, language is often characterized by circumlocutions, ambiguities, ellipses, and repetitions, which are among the many vagaries in human communications. The basic premise of SPA is that natural language systems, instead of seeking to "train" humans to accommodate their speaking patterns to the speech interface, must be able to adapt to the less than perfect speech produced by humans.

Can you provide more technical details? How does one use your research?

I've designed a BNF (Backus-Naur Form) table consisting of 70 sequence packages, which are a series of related turns and parts of turns discretely packaged as a sequence of (conversational) interaction. The parsing structures contained in each sequence package consist of a set of non-terminals—context-free grammatical units and their related prosodic features—for which there is a corresponding list of *interchangeable* terminals: words, phrases, or a whole utterance. What distinguishes the SPA-designed BNF table from a conventional table is that its parsing structures are not syntactic components, encompassing parts of speech and phrases, such as N, V, ADJ, NP, VP or ADJP. Instead, they are *sequentially-implicative* units, meaning that their formal grammatical representation is defined by sequence as opposed to syntax.

Is your system grammar-based or statistical?

Both. It is semantic grammar-based, for those clearly defined sequence packages that contain specifically marked boundaries and specifying package properties, and statistical, using N-grams to depict the probabilistic occurrence of a sequence package structure when one is not so clearly defined. I will say that since sequence packages are both domain-independent and language-independent, the costs of using a statistical approach are not prohibitive as they are for those applications where data changes dynamically, as is the case of seasonal applications.

For further information:

Linguistic Technology Systems	Fort Lee, NJ	Linguistics and speech consulting	(201)224-5096; www.lingtechsys.com