What if we used poetry to teach computers to speak better?

November 18 2010

A better understanding of how we use acoustic cues to stress new information and put old information in the background may help computer programmers produce more realistic-sounding speech.

Dr. Michael Wagner, a researcher in McGill's Department of Linguistics, has compared the way French- and English-speakers evaluate poetry, as a way of finding evidence for a systematic difference in how the two languages use these cues. "Voice synthesis has become quite impressive in terms of the pronunciation of individual words," Wagner explained. "But when a computer 'speaks,' whole sentences still sound artificial because of the complicated way we put emphasis on parts of them, depending on context and what we want to get across."

A first step to understanding this complexity is to gain better knowledge of how we decide where to put emphasis. This is where poetry comes into play. Wagner has looked at prosody, which means the rhythm, stress and intonation of speech. Poetry relies heavily on prosody, and by making a comparison between languages, he is able to uncover how prosody functions differently in English and French.

Working with Katherine McCurdy at Harvard University, Wagner recently published research that examined the use of identical rhymes in each language. "These are rhymes in which the stressed syllables do not just rhyme, but are identical, such as write/right or attire/retire," Wagner explained. "It is commonly used in French poetry, while in English poetry it is considered to be unconventional and even unacceptable."

Wagner gave the following example from a book by John Hollander:

The weakest way in which two words can chime

Is with the most expected kind of rhyme —

(If it's the only rhyme that you can write,

A homophone will never sound quite right.)

The study shows that identical rhymes fit into a general pattern that also applies outside of poetry: even when repeated words differ in meaning and merely sound the same, the repeated information should be acoustically reduced, otherwise it sounds distinctly odd. "It's sort of a bug of the way English uses prosody," Wagner said, "but one that hardly ever creates a problem, because it occurs so rarely in natural speech." Wagner is now working on a model that makes better predictions about where emphasis should fall in a sentence given the discourse context. His findings were published in the journal *Cognition* and received funding Quebec's Fonds de recherche sur la société et la culture and a Canada Research Chair in Speech and Language Processing grant.