

Computer Translations Lose Too Much Information When Used In Adversarial Situations

Preventing terrorism is universally considered an extremely important undertaking. One of the most difficult problems facing American agencies that deal with this situation is the language barrier. Many terrorist organizations that wish to target America have members that primarily speak Arabic. Unfortunately, machine translation cannot help in this task.

Investigations of the September 11th attacks in New York City and Washington DC led to information known to American intelligence agencies which hinted at the attack. This information was gathered on September 10th of 2001, but was not translated until two days later. In general, information on suspected terrorists is gathered at a rate much faster than the rate intelligence agencies are able to analyze it, resulting in much information not being examined until it is too late. As an example, Michael Erard writes that “every three hours, NSA satellites sweep up enough information to fill the Library of Congress” [Erard 2004, p. 56].

Many believe that it is infeasible for humans to undertake this task alone, and the aid of computers is necessary. Several large-scale translation devices are either being used, or being developed. Others, however, are skeptical, and don't believe that a computer can replace a human. Much in a language can be gained by tone of voice, pauses in speech, cultural references, and many other subtleties that are lost when a computer translates a phone call or other conversation.

As an example, other American government agencies have had problems when listening to English conversations. The Federal Bureau of Investigation often has wiretaps on those suspected of organized crime, and yet often cannot decrypt the information received. The phrase, “I want you to talk to him” can be purely innocent, but adding a slight pause before the word “talk” can drastically change the meaning of the entire sentence. People would notice the subtlety, while computers cannot, and this is without any language transition at all.

Another example can be constructed to use American culture, which the average Al Qaeda member would most likely not understand. The phrase, “tomorrow, the Red Sox win it all”, can be spoken between two die-hard baseball fans, or by two terrorists making an allusion to the underdog, that has lost for a very long time, finally achieving victory. A fundamentalist culture, such as that shared by Al Qaeda, must have plenty of such references that would have equivalent levels of ambiguity when translated to English, and read by someone who does not know Islamic culture.

While America does have a terrorist problem, which stems primarily from Islamic terrorists, there are far more people who know the Islamic culture and Arabic language

and are not terrorists. Instead of translating Arabic to English for non-Arab Americans to interpret, more Arab-Americans should be found to analyze the original Arabic. If there is still too much information to be analyzed, then perhaps the American security agencies are making their surveillance much too broad, and should therefore narrow the scope instead.

References

Erard, Michael. "Translation in the Age of Terror." *Technology Review* (March 2004): 54-60.

Hutchins, W. John. *Machine Translation: Past, Present, Future* (Chichester, UK: Ellis Horwood, 1986), chap. 2, "Precursors and pioneers"; chap. 8, "Expectations and criticisms"; chap. 19, "Present developments and some future prospects."

Weizenbaum, Joseph. "The Computer and Natural Language." Chap. 7 in *Computer Power and Human Reason* (New York: Freeman, 1976), pp. 182-201.