## (U) Strangers in a Strange Land?

FROM: (C) ▮▮▮▮▮▮▮▮▮▮ MENA, S2E21 and
(U//FOUO) Anita H. Kulman, Human Language Technology (S23)
Run Date: 10/05/2006

*(U//FOUO)... not so much: How Language Analysts are supporting Human Language Technology research.*

(U//FOUO) Once past the cipher lock on the door, at first glance, the office does not seem very different from offices in Analysis and Production in the Big Four. But, looking more closely, it's not the same. *More desk space? Wider aisles?* Maybe. *More equipment?* Yes, there seem to be more computers. Many desks have two, and the displays are strange: software tools not seen in operations, funny windows with interesting graphics, scatter plots of dots and circles, squiggles, shades of color.

(U//FOUO) *And the people working here?* They are mostly young men who look very smart and intense. (There are women, but their desks happen to be in the back so you see them only once you find your way through the cubicle maze.) No one seems to be talking because all are focusing fixedly on their work. Behind each of them, at arm's reach, there are white boards. In fact, there are whiteboards everywhere, and they are covered with strange words and with equations.

(U//FOUO) Often at the larger ones, two or three people stand with their noses almost pressed to the board, writing equations and discussing the work, and using language that we don't hear in S2, but it's not a foreign language. *What are they talking about?* Gaussian Mixture Models, Hidden Markhov Models, DET plots, EERs, Segmenter, OVENCAPTIVE....

(U//FOUO) *Where are we?* In the basement of the **Research and Engineering Building, in R64, the Human Language Technology Research Group spaces.** And who are these people? They are HLT researchers, mainly electrical engineers, mathematicians, and computer scientists who are working to develop the algorithms that will be the foundation for tools to assist SID operational analysts process multilingual SIGINT intercept. Here, work is done mostly on **speech** projects, but close-by, in other R6 spaces, similar HLT work is being accomplished on **text** and **image** intercept. They are researching ways that computers can do some of the lower-level analytic tasks so that analysts will be free to do the upper-level work.

(U//FOUO) *And who are the strangers here?* We are Language Analysts from S2, who have left our desks in operations to come here for a few hours every week or for longer details to contribute precious language and analysis expertise to this mix of research specialties. Here we work as Scientific or Computational Linguists and Language Technologists.

(U//FOUO) At first, it's hard not to be intimidated by the researchers' exceptional technical expertise, but language expertise is desperately needed for research efforts such as Language Identification and Speech-to-Text research, as well as on many other projects. The researchers may be intimidated by our exceptional technical expertise in language analysis. They are happy to welcome us, and our work is highly appreciated.

(U//FOUO) *And exactly what kind of work do we do here?* It's real language work. We listen to volumes of SIGINT intercept and label the files according to various characteristics. This can be as "simple" as labeling what language or dialect is being spoken-- which we know is not so simple since there are some 6800 human languages to choose from. Or it can mean transcribing the messages very, very carefully.

(U//FOUO) The researchers give us tools which align our marks with the voice stream of the message. In this way, we help to create large corpora, databases of language material that the researchers later use to train the algorithms they create to recognize words or phrases and to write them out in the foreign language native orthography, or to distinguish various languages from one another.

(U//FOUO) For S2 Language Analysts who volunteer a few hours per week or who take longer diversification tours to assist R64, the Human Language Technology Research Group, the environment and the language work itself are very different from what they do in their operational assignments. The environment is more relaxed and more intense at the same time. That is, the frenzy of direct mission-related work is absent, but the requirement to do the work very carefully and to be incredibly precise is exacting.

(U//FOUO) In the end, the experience that SID Language Analysts have working in R6 is very rewarding. Their expertise is put to use to help other language analysts, so they become real force multipliers. They become tutors in language skills for the research engineers who do their own jobs better once they better understand language skills. The Language Analysts extend their own expertise by learning how to use sophisticated tools, and even more, they help the researchers to significantly improve those tools. Everyone benefits.

(U//FOUO) If you think you would like this kind of experience, email or call Anita Kulman ████████████████ to check it out.

**"(U//FOUO) SIDtoday articles may not be republished or reposted outside NSANet without the consent of S0121 (DL sid_comms)."**