DARPATech, DARPA's 25th Systems and Technology Symposium
August 8, 2007
Anaheim, California
Teleprompter Script for Dr. Joseph Olive, Program Manager, Information Processing Technology Office

Language Translation

» **JOSEPH OLIVE:**
If you were to walk by a wall like this one

...It would mean nothing to you.

And later, you might not understand why this had happened.

But imagine a future where our military personnel are equipped like this

Today, many of our military personnel embedded in foreign environments are unable to understand the language of their surroundings,
unable to communicate with local populations and unable to absorb mass media.

In the Arabic speaking world alone there are hundreds of
radio stations,
television stations, newswires, and websites producing thousands of hours of audio and millions of words of text every day.

For effective intelligence gathering in support of combat missions, foreign language information
must be translated, and distilled to gather tiny bits of relevant information about groups, individuals, and localities of high interest to

our military.

In the case of urgent, mission-critical information,
a delay of just a few hours may be too long.

The supply of qualified linguists is always less than the demand.

And it is impossible to predict what new language skills will be needed in the future.

DARPA has the vision to solve the language barrier problem.

Our vision is to create a universal translator,
a technology that would allow us to communicate with people with whom we do not share a common language and enable our soldiers to have situational awareness whether it is through understanding of local media or their surroundings.

To better understand why fulfilling this vision is so difficult we have to examine the nature of human language.

The fact that very young children acquire language so easily
belies its complexity.

Because language is so natural to us, most of us are not even aware that it is actually an abstract code for conveying information and concepts.

This code is made up of sounds that form words, and sequences of words that form messages.

Language, however,
is far from an exact code.

In fact, it can be quite ambiguous.

Individual words often have many meanings.

Sound sequences vary because of differences among individual voices, accents, and speaking styles.

Environmental factors, such as noise and echoes, can also affect communication.

Humans use a variety of mechanisms to understand the messages that language conveys.

We have an amazing capacity to quickly adapt to sound variations, background noises, differences in word usage, accents, and new words, just to name a few.

We also combine knowledge of a language's structure with an understanding of the context of the message.

Let me show you an example of how important context can be for understanding speech.

When I say
"Machines may wreck a nice beach,"

Given the context of my talk, you probably thought that I said this…

But actually, I said

If you didn’t understand me, it was because you assumed that I was talking about speech.

Now, what about written language?

Words can be represented in writing,
but the written text cannot convey the intonation and timing differences that are so important in spoken communication.

And of course the problem is compounded when trying to read handwritten material.

The additional information that could come from speech is still missing, and we are also confronted with the uncertainty of deciphering poorly formed text.

Despite all these difficulties,
humans are very good at communicating through both spoken and written language.

But what about the use of human language by computers?

In science fiction,
we find many examples of computers like HAL in *2001*
and C3PO, in *Star Wars*, that have full communication skills and translation capabilities.

But then again…
it is science fiction.

In reality, computers today do not have the knowledge, experience, deductive reasoning, or language comprehension needed to **understand** speech and text.

Fortunately though, DARPA has long understood the importance of computer language processing to the military and has invested substantially over the past three decades in language processing research.

In the beginning,
because of our limited understanding of the science and limited computer capabilities, speech research was restricted to recognizing a few words by a single talker.

Over time, computer vocabularies increased and speech recognition systems became
speaker-independent.

Some of this research has migrated to the commercial sector producing applications such as
automated call centers, travel reservations services and
dictation machines.

But recent advances in DARPA sponsored research have produced technologies with extreme value to the military.

One of these technologies is a two-way translation device, developed through DARPA's Translation Systems for Tactical Use program, which is being used today in Iraq.

This device enables service personnel to talk in natural English and have their speech translated automatically into Iraqi Arabic –
and to hear spoken Iraqi responses translated into English.

The device is designed for specific dialogue types,
so its language repertoire is limited, but it is effective and highly valued

by our troops.

Here is a scenario in using this technology today.

Last year we began a very aggressive program called Global Autonomous Language Exploitation,
or GALE that is developing technology to:

Translate foreign speech and text into English text and,

Distill and extract precise, actionable information...

...and to do all of this accurately and quickly.

This program is extremely challenging.

Translation, for example, cannot be accomplished by simply looking in bilingual dictionaries and arranging the words according to grammar rules.

Words have multiple meanings as well as multiple grammatical functions.

50% of English verbs have more than one meaning, and often the same word can be a noun or a verb.

At present, DARPA machine translation systems work as follows:

Using large translated corpora, the computer finds short segments in the text of the foreign language and retrieves corresponding segments in the English.

The most probable segments are joined in the most probable order to generate the translated English sentences.

This method,
called statistical machine translation, has speeded progress and is currently providing adequate translations for certain applications.

One application is a system that monitors open source media,
such as TV and radio transmissions and news articles on the Web.

As you can see,
the system is able to find information in foreign material that does not exist in English sources.

This system is currently far from perfect, but,
it is good enough to extract basic information from an article or broadcast and allow analysts to decide whether they'd like it translated more thoroughly by humans, thus reducing the flood of Arabic media down to a manageable volume.

For example, United States Central Command receives about 5000 Arabic language articles per week.

After translating and filtering these automatically,
CENTCOM sends only about 300 to human translators.

This is a 94% reduction
in human labor.

As translation quality improves, we expect to reduce or even completely eliminate the need for human translators.

To improve automatic translation, GALE is integrating transcription, translation, and distillation.

Sharing information among these processes will provide more alternatives to be considered for a more accurate translation.

The choice of proper synonyms will improve after proper topic identification.

But the major improvements will come as the translation process learns to use extensive grammatical analysis, knowledge of
word functions and
word properties
to prevent the computer from producing sentences like this,
where the computer apparently confused the subject and object …

When the original Arabic actually said this.

Before GALE, computer translation accuracy was about 55% on structured text and 35% on structured speech.

After a year of GALE program research,
our accuracy for Arabic structured text climbed to 75% on 90% of the documents tested.

For speech, our accuracy increased to 65% on 80% of the segments tested.

However, these results are still far from our ultimate goal of
95% accuracy
on 95% of text documents and 90% of speech segments.

As you can see we have come a long way from the days of automatic recognition of a few isolated words.

We have also produced technologies that are helping the Department of Defense.

But despite our great progress, we still have a long way to go to achieve our vision of universal translation.

There are obstacles at every point in language processing.

One major difficulty,
even for humans,
is to understand speech in noisy environments.

There is a similar problem with documents, since they may be stained, torn, or badly copied or transmitted.

Although this may seem like a relatively minor problem, interference poses a particularly difficult challenge for computers because they lack the cognitive functions that humans use to fill in gaps from context and from comprehension.

Overcoming such degradations is essential for using translation technology in the street or on the battlefield.

Another major challenge is that current automatic translation systems are very specialized.

A computer translation system does well when presented with topics it is programmed for, but its accuracy decreases when faced with topics outside of its specialization.

We are currently exploring many means of addressing this problem.

One of them is to research dialogue strategies in order to implement the human ability to interactively ask for clarification.

In addition, we need to develop methods to make language technology quickly adaptable to new languages.

We are already studying translation of Arabic dialects, but this problem pales in comparison to the difficulty of dealing with languages that have entirely different vocabularies and grammatical structures.

To achieve our ultimate vision, we need to develop computers that understand more than just grammar and word functions.

We need to develop computers that learn the meaning behind the language.

While working towards our ultimate vision of
universal translation,
we will integrate new technologies as soon as they are developed
to ensure that our warfighters have the best and most current
technologies available.

The men and women in uniform need these technologies sooner, rather than later.

But we have big challenges and we need your support.

More important, our men and women in uniform need your support.

Next, you'll hear from Sean O'Brien about Computational Social Sciences.