



Restricted language resolves ambiguity:

In the upper balloon, the argot-speaking police officer is telling headquarters that "a wardrobe with a profound voice hit a truant with hair in brush.

The ladle-like truant is as round as a little pea."

In the lower, balloon, he is saying in unambiguous, restricted police-speak that "a heavily built man with a deep voice hit a hooligan with a crewcut. The hooligan is as drunk as a lord and suspicious-looking."

There are times and places when plain speaking saves human lives. Soon restricted languages may also be helping to replace humans, by automating many word-based information exchanges.

LESS IS MORE

Trevor Jackson

When an air traffic controller is asked by his toddler, "Where's my teddy, Daddy?" he doesn't reply, "Maintain your course, position zero five three, from the dining room table." He doesn't confuse "Air-traffic speak" with "Homespeak."

In fact, it is the need for precise communication in specific environments that has triggered both the creation and the use of such "restricted languages" as Airspeak and Seaspeak, which Edward Johnson and his colleagues at Cambridge (UK) University's Wolfson College have developed.

"My experience with mariners, pilots and air traffic controllers is that when they walk into the operational situation, they enter a mental set," Johnson explains. "The 'operational language' takes over so that they click into a special way of behaving." Further, these "operational communicators" become very perceptive of the dangers of natural language.

Communications such as air traffic control or ship-to-shore, or – in the future – between the French and English police at opposite ends of the Channel Tunnel, require precise exchanges of information that directly affect human safety. But as Johnson notes, "Natural language is not at all precise. That is why I like working in restricted language. Only one interpretation is possible."

Speakers of a restricted language agree that whenever they enter their "operational situation," they will obey a set of rules. These rules ensure that communication is unambiguous; that there is one, and only one, interpretation for any utterance. Formally clearing away the ambiguities – the "noise" of natural language – in advance of the actual communication reduces errors in understanding and paves the way for automating many word-based information exchanges.

Restricting natural language to a core of well understood elements is not only enhancing public safety in international transportation systems. It is also influencing the development of viable machine translation (MT) and speech recognition systems, and may permit the kind of smooth communications required by a united Europe.

"ALLO, ALLO, ALLO"

Johnson and his colleagues at Wolfson College are now developing Policespeak in response to a request from the Kent Police. The force wants a simplified language, stripped of the ambiguities of natural language and police jargon, to enable fast and accurate communication with their French counterparts when the Channel Tunnel opens in 1993.

As Superintendent David Lambert, chairman of the Kent Police Committee considering safety issues raised by the tunnel, has said: "The day we're faced with a Hillsborough or a King's Cross disaster (recent public disasters in the UK which highlighted communication problems between rescuers), we will need very close cooperation with the French in the heat of the moment."

Johnson explains: "Policespeak is not one language, it's a pair of languages in concord. The English of policing and the French of policing are separate languages. The idea is to bring them closer together, so that they are interchangeable. Policespeak is designed to resolve a human problem of communication.

"Here you've got two police forces that will have to resolve life-threatening situations in the tunnel frontier environment. It's merely a case of saying: 'Before we start this, let's get our terminology set.' For example, we can agree to ban certain words because they are just too ambiguous, and that ambiguity can be very dangerous. If we're talking over a VHF radio, we can both agree not to convey meaning by intonation. It's like a game, we agree to the rules before we start." As soon as you pick up the radio you are declaring your context – that is, the highly specialized "operational situation" of Policespeak.

Early in the 80s, it occurred to Johnson that research in MT ought to concentrate on areas where there was prior agreement about the language being used. "This would clear away a whole lot of garbage and ambiguity – which is the problem that machine translation has had in the past."

"Translation is most dependable, whether a machine does it or a human being does it, when the context is fully understood and the communicative purposes of the speakers are well known. The more restricted you get, the more reliable is your translation."

Johnson and his team, collaborating with British Telecom, have also been developing an interactive MT system for restricted languages. The system, known as LINITEXT (Linguistic INterpreter for International Text), is being developed specifically for business communications. All the research has been guided by principles that emphasize the nature of restricted languages. For example, researchers closely control the lexicon and syntax, as well as specifying the domain and language pair. They rely primarily on corpus analysis, and have no ambition to process arbitrary text.

For the Policespeak project, Johnson and his team are currently analyzing a corpus of 50,000 words from police communications. Expressions are identified according to their communicative function, including the locale of the communication, the particular operation the speaker is discussing, and the speaker's overall purpose in saying it. This analysis is used to determine exactly what English is essential for conveying particular messages in the operational situation.

Two other restricted languages, Airspeak and Seaspeak, are the prototypes for Policespeak. Johnson has been involved in the development of both.

PLANE ENGLISH

Airspeak, an English-based radiotelephony (RT) system, is the restricted language of air traffic control that has taken shape over the last

40 years. It has evolved through a process of commonsense restriction and contraction of natural language. Pilots and air traffic controllers who actually use the language have instigated and orchestrated its development.

Airspeak is probably the world's most successful restricted language. But critics claim the increasing volume of airtraffic worldwide demands more than RT: they would like to see speech-independent radar-based methods of communication or data links. Johnson has recognized these criticisms in the foreword to

Fiona Robertson's Airspeak guide: "The odds are heavily stacked against fail-safe RT communications. Indeed with physical impediments such as blocked frequencies and simultaneous transmissions which occasionally inhibit radio contact altogether, propagation noise, background interference, electrostatic noise, the far from ideal acoustic environment of the flight deck, all of which

contribute to the degradation of the signals received by the brain, it is astonishing that RT is as effective as it is."

But these limitations and criticisms focus on the medium that carries the message, rather than on the restricted language itself. Real time speech continues to be used, despite these problems, because it does not compete with the visual communication systems already in place in the cockpit and air traffic control tower. Moreover, Johnson observes that the delivery technology of restricted languages is important and can affect the language developed. For example, in a simplex radio system, the speaker can't be interrupted. This means you can't be helped or reassured by the listener who might give you verbal clues that s/he understands the message and can fill in the blanks.

In this type of system you've got to get it right the first time, which means you're going to form the message quite carefully before you begin. In this case, the restrictions imposed by the delivery system actually help perfect the communication.

SHIVER ME TIMBERS, IT'S SEASPEAK

In 1987, the International Maritime Organization approved this distilled (and in some ways enhanced) version of English as the lingua franca of maritime VHF communica-

tion between ships, ship-to-shore, and vice versa. It was the maritime professionals themselves who had demanded a simple, unambiguous language for their communications, and now many nautical colleges teach Seaspeak.

In addition to the English language, Seaspeak brings together several elements, including the special language of maritime affairs, restraints imposed by the medium of transmission, and the functions of maritime communications. The language has three basic stages: making contact, exchanging messages, and terminating contact. In determining the content of the messages, Johnson realized that a vast set of standard phrases was clearly unworkable and that a fluent knowledge of the English language should be unnecessary.

The developers found they couldn't simply adapt Airspeak for mariners because the operations of the two domains are quite different. As Johnson explains, "Air conversations are highly predictable due to the uniformity of airports and operational procedures. In that environment, a 'slot and fill' phraseology serves very well." That is, a particular operation remains constant, while the values, names, codes etc., change according to local circumstances. But standardizing a complete harbor entry call is quite another matter. There is a huge diversity of harbor types, traffic systems, and ship types, as well as additional complicating factors such as pilotage, tug services, the presence of pleasure craft, sea bed variations, and tides.

Johnson says: "It was necessary to adopt a research strategy that would allow a more flexible use of language. This was done by examining the language of maritime VHF communication in terms of the mariner's reason for calling, rather than to focus on categories of operational situations."

In carrying out the research, the major sources of data analyzed by Johnson's team were about 1,600 maritime VHF transmissions, published works containing recommendations and phraseology for maritime RT, nautical dictionaries and almanacs, specialist communications publications, a set of 200 logged VHF conversations provided by shipping companies, and information from professional mariners.

An IBM 370 mainframe computer was used to analyze this large body of information. Once the contents of messages in the data had been examined, it was then possible to work out the simplest and most straightforward ways of transmitting individual items. Since means of transmitting course, speed, time and position have evolved over the years, it was necessary to create a standard that was in line with current regulations and

practice. In this example, "My ETA is one six zero," the numbers clearly refer to time. But the meaning is less clear in the statement, "The wreck at one six zero zero is one six miles south of Star Point," particularly to a non-native speaker of English, since the meaning depends heavily on the use of the words 'at' and 'is.'"

It was therefore proposed that times be prefixed by the word "time," position by the word "position" and so on. This process of "apposition" was not new, but the proposal that it be adopted as standard usage in maritime VHF was.

Overall, Johnson notes, "The challenge has been to decide just what is necessary in terms of grammatical constructions and words in order to define that sub-set of the English language that is adequate for maritime VHF communication."

RESTRICTION IMPROVES RECOGNITION

Restricting natural language improves the accuracy of speech recognition systems, in part because it reduces the recognition problem. As Jim Baker, co-founder of Dragon Systems puts it, "If you know in advance the specific vocabulary the speakers will use, you can substantially reduce the errors in recognition by a factor of four."

For example, while it has always been the company's goal to develop a general purpose speech recognition system, the development of earlier, more restricted systems such as Voice Scribe, have directly aided the development of Dragon Dictate.

Voice Scribes 400 and 1000 had recognition vocabularies of 400 and 1000 words respectively and were used by clients in specific application areas, such as radiology and manufacturing. The new system, Dragon Dictate, recognizes 30,000 words and is considered speaker- and application-independent.

Baker maintains, "Anyone who is otherwise pushing the limits of the state of the art of speech recognition can get definite improvement by developing an application-specific system."

Jim Moser of Kurzweil Application Intelligence agrees that focusing on a specific application area significantly reduces the recognition problem, and allows the developers to concentrate on making the system act on what it has understood.

That strategy has resulted in products for three specialized medical environments: radiology labs, emergency rooms, and pathology labs. Coroners can now produce their final autopsy reports while they are performing the autopsy. They merely glance at the screen while they are working and ver-

"I'm interested in the shared meaning developed between two individuals, rather than the link between the language I use and the thought processes that go on in my brain. That, in a sense, is too personal."

— Edward Johnson

RESTRICTED LANGUAGE: A Brief History

Early in the Renaissance, traders developed and used the original lingua franca to do business all along the Mediterranean. This mixture of Italian, Spanish, French, Greek and Arabic evolved to support a specialized activity, much the way AirSpeak has been developed by air traffic controllers and pilots in recent years.

At the same time that merchants were speaking MerchantSpeak down on the docks, Francis Bacon and other scholars were trying to come up with a representational system. This language would be used to transmit scientific thought unimpeded by the ambiguities of natural language. Their purpose was broader than the merchants' and the method was scientific.

Other attempts at creating artificial languages have been no less ambitious. During the French Revolution, people dreamed of uniting humanity through a common language. The 18th century mathematician Marie Jean Antoine Nicolas Caritat, Marquis de Condorcet, believed that only a rationally based universal language, independent of the ravages of vernacular evolution, would be able to preserve and transmit human knowledge. Condorcet, together with Delormel, outlined an artificial language and presented it to the National Convention in Paris in 1796. Their work led directly to Maimieux's Pasigraphy (1802) – the first completely artificial language ever to be taught.

Maimieux's work was largely based on a comparative study of Indo-European languages in an attempt to express the features of a universal grammar. But even in the fervor of the Revolution, researchers were wary of the losses

imposed on users of restricted languages. Johnson writes: "They were concerned that such inflexible systems could impede scientific enquiry. They could not sanction as democratic any system which might remain inaccessible to the masses and they could not reconcile themselves to the potential suppression of imaginative thought and language, which the adoption of such schemes inevitably implies."

Equally humanitarian motives led the British linguists Ogden and Richards to develop Basic English in 1930. They wished to limit a living language in order to develop an international second language. Paradoxically, the Allied Forces were the first to test this humanitarian endeavor. They found it wanting and abandoned it.

Basic English has been the basis of Caterpillar Fundamental English, designed by Caterpillar Tractors for use in their technical manuals. It has successfully been adapted for use in other constrained environments, such as the aerospace manufacturing and computer industries.

The motive for designing these more recent adaptations of basic English are far less ambitious than Ogden's original purpose. Our contemporaries only seek precise communication within a well-defined environment.

Though we value the dream of Ogden and Condorcet, we benefit from the more reliable goals of researchers like Edward Johnson. For while hundreds of artificial languages, including Esperanto, have been created, none have been adopted for the widespread use their creators intended. On the other hand, scores of restricted languages or operational sublanguages, which have either evolved naturally or have been purposefully derived, are actively in use today.

bally respond to the prompts. A few words can actually trigger entire paragraphs to be incorporated into the report. For example, a doctor might say, "no conditions in the liver," and the system generates all the sentences that describe the specific abnormalities understood by "no abnormalities."

Restricting natural language may simply mean recognizing the essential components of an existing domain-dependent language – rather as we talk about "legalese" or "computerese." Even this informal degree of restriction can increase the effectiveness of speech recognition systems.

According to Deana Murchison of Speech Systems Inc., "All the projects we undertake begin with 'naturally restricted' languages." These are the "languages" already in use in the air traffic control tower, at the bank, and in specialized medical laboratories. Speech Systems Inc. has a contract with the US Department of Transportation to develop a training system for air traffic controllers. The system includes a grammar, jointly devised by Speech Systems Inc., a Tarzan, CA company, and the Dept. of Transportation, which allows the speech of the trainees to be recognized by the system and to be responded to by a synthesized pilot's feedback.

Speech Systems has developed an expert system that parses the trainees' utterances to determine their meaning, then simulates the aircraft movements specified by the trainee's speech on a radar screen. The system may be

recommended to training centers throughout the US, effectively training new controllers in the restricted language recognized by the system.

POST-1992: RESTRICTED EURO-BABBLE?

Edward Johnson told a conference in Paris last year that the restricted languages of Airspeak and Seaspeak were not simply isolated instances of "purpose-built languages." In a paper entitled *The Role of Languages in Economic Competition*, he argued that, "it is probable that we will see more initiatives of this sort in Europe in other specific areas of human endeavor which up to now have used natural language."

He explained that purpose-built languages were already growing in areas like banking, stockbroking and international freight forwarding. "All these subdivisions are already able to communicate," he said. "They have developed means of communication which circumvent the language barriers. It's something that human beings are extremely good at. There's so much concentration on language barriers, as though attachment to one community was primary," when in fact, "the barrier of the EEC after 1992 will be that outer frame of all the countries it contains." For example, once the British police begin to depend on the Spanish police to ensure that South American coke barons are not entering the EEC in order to reach British nostrils, the Spanish and Brit-

ish police must communicate efficiently.

"You can't say that we have to learn all the languages of the EEC – there are 72 language pairs! It's beyond the capability of human beings to do so," Johnson said. Instead, he urged wider use of restricted languages. If that occurred, "there would be massive circumvention of language barriers which will involve a switch from language centrality to purpose centrality." The result he believes, would be systems of communication exhibiting operational efficiency approaching that of air traffic control.

One might argue that restricting our vocabulary, syntax or grammar would affect not only the way we use natural language in unconstrained situations, but also the way we actually think. Will the imposition of these highly formalized, purpose-built languages, even though they are extracted from human experience, ultimately reduce the possibilities for new experience?

Johnson steps gracefully and purposefully away from this quagmire, when he says, "I'm interested in the shared meaning developed between two individuals, rather than the link between the language I use and the thought processes that go on in my brain. That, in a sense, is too personal. All our research here concerns developing an understanding between two human beings."

Trevor Jackson is a London-based journalist specializing in human interest anti-Poll Tax stories.