Space Languages

Dan Hays
Psychology Department/Cognitive Systems Laboratory
The University of Alabama in Huntsville 35899

Abstract. Applications of linguistic principles to potential problems of human and machine communication in Space settings are discussed. Variations in language among speakers of different backgrounds, and change in language forms resulting from new experiences or reduced contact with other groups need to be considered in the design of intelligent machine systems.

The languages that people and machines will use in near-Space settings will be about the same as the ones that they use on Earth, at least in the short run. These languages, and their users, will operate according to exactly the same principles of linguistic structure and change, whatever the setting.

So far as language use by humans interacting with other humans, for the most part we do not have to worry about it for missions in the near future, whether the humans are working in Space or are members of ground organizations which interact with orbital or planetary locales. People are such capable users of language, and the kind of language we have as a species is so functional, that it need not concern us so much as, say, the reliability of energy transduction and distribution (which is currently seen as critical and somewhat problematic) or the psychological and social effects on Space residents of isolation, restricted quarters, and scheduling pressures (which still do not appear to concern the U.S. space effort as much as they should).

Even so, our complex and in many ways delicately tuned language faculty is a *very important human sub-system,* so to speak. For this reason alone it may bear some examination as we make plans for Space residence. Other more concrete reasons include the following:
  • Several human languages will be spoken in near-Space, in some cases in the same locale. Even when the speakers are bilingual, problems can arise when assumptions differ, owing to differences in the native language cultures.
  • Within the same basic language, different occupational or regional dialects may be spoken, allowing not just misunderstandings of terminology or nuance, but also invoking potentially troublesome group identity conflicts.
  • Errors and misinterpretations occur occasionally because of economy of expression and the partial, general nature of reference in human language.
  • Yet other errors of communication, either between people, or in data entry or video-screen perception, seem to occur because of lapses of attention or other "low-level" processing problems.

When it comes to humans interacting in information-rich ways with machines, or even sufficiently advanced machines interacting with other machines, linguistic issues are much more problematic. Since most foreseeable Space habitats will be heavily mechanized, and may be fairly heavily automated, characteristics of human language, and indeed of languages suitable to machine interchange, should concern the designers of complex devices and computer-based information systems. This is true whether these devices are "intelligent" in the usual senses of reasoning logically or heuristically, using analogies, profiting from experience, and so on, or

PRECEDING PAGE BLANK NOT FILMED

just are *sensitive to situations and interactions.*

Communication using some subset of human language has been of concern to computer system designers, perhaps especially during this decade. Diagrammatic and other graphic display of information is more frequently used, also, as machine capabilities increase, and as computer professionals become more concerned for human interfacing (compare, for example, computer displays in the first SpaceLabs with suggestions for the next U. S. space station). As time goes on, we may expect:
* more *actual language analysis,* both of human input, and in generating output from computers, rather than just recognition and delivery of set phrases with slots for variable information,
* developments in *graphic and pictorial* information exchange,
* confrontation and exploration of issues of *semantic analysis* that is rich enough and flexible enough to mimic human understandings, impacting the structure, for example, of content-linked databases or knowledge bases,
* more explicit appreciation of *'pragmatic' and other socially relevant levels* or aspects of language usage.

Advances have been made in robotic devices. But we have hardly explored either their
* social capabilities [4] [8] or symbolic conventions for the informational exchanges that should occur between humans and devices, or among autonomous action devices [7].

Though it was suggested above that "space languages" will not be much different from ordinary Earth languages, really distinctive communication conventions might either be designed or evolve.
* Signaling conventions might evolve or be designed for special situations in extraterrestrial work. For example, gestural or iconic signs useful for work outside of spacecraft seem to exist already and should undergo development. Location markers, whether visual or in some other medium, may be conventionalized for autonomous devices and space craft.
* It is possible that certain conventions might be useful for interchanges among humans that speak different Earth languages.
* If remote sites are inhabited by people, they will almost certainly diverge somewhat in language terminology and possibly in language form as time goes by, just as languages on Earth change.
* Since the beginning of near-Space exploration by people, distinctive jargon has developed (some of it widely known) among Space workers.

The reality of language change will inevitably concern persons who design or work with computing devices where knowledge is stored and communicated via ordinary language. At the very least, new terminology and descriptions of new objects must be incorporated into knowledge-rich systems; and their data structures and interpretation routines must be flexible enough to allow for variation and approximation.

Language Variation

Linguists say that human language varies *synchronically,* at a given time, when we note hundreds of distinct languages and many more dialects; or *diachronically,* over periods of time. References to historical linguistics include [1] and [9].

We are familiar with the phenomenon of new terms coming into common use, and the dropping out of others. A number of processes can be involved. A term might be borrowed from another language ("détente"), or made up from existing parts ("television"). Metonymy, or referring to something by the name of something closely associated, is another ("a Bordeaux" for a wine made in that region, or "a Winchester" for a kind of disk drive or rifle). Other kinds of variation include shifts in meaning, where a term comes to have a related or even an opposite sense; shifts in syntax, for example, in the present instability of adverbial suffixes in English; change in performance style, and so on.

These changes take place against a background of great concurrent stability in a language, which must be predictable for communication. Nevertheless, and sometimes to the dismay of academics and newspaper columnists, language changes at all levels. Some changes are certainly prompted by situational novelty, or contact with others who speak differently; others seem almost intrinsic to language as a complex system. For example, some tendency to simplification of expression, at least for matters of frequent usage, may relate to a basic impulse of cognitive simplification. Exact kinds of change are difficult to predict. However, others follow precedent or cultural custom. For example, both English and Japanese have borrowed technical terms from other languages, though in somewhat different ways. Chinese languages have more frequently constructed new terms by metaphorical extensions of native morphemes.

The dynamic nature of language, together with the fact that Space workers will be confronting situations that Earth planners might not be able to envision in detail, argue for setting up computer-based information systems with some flexibility for growth and modification.

A potential source of tension in future machine knowledge technology in Space is engendered in the contrast between the changeable nature of human language, and the predilection of machine systems, and probably 'rational' systems in general, for fixity of reference. Put another way, any adequate means of representing "meaning" for machine systems must also handle "change in meaning".

Having expanded or modified meanings in machine processing implies that all concerned parties, whether human or electronic, must have a way of learning about the extensions or adjustments that are relevant to their activity. This would apply to occasionally interacting knowledge bases within one setting, such as a space station or small ground facility; and it would be of major importance to autonomous mobile devices, which must confront and symbolize novelty frequently, and which should be able to relay some of this information to companion devices and to people.

If outposts or settlements of humans come to exist in remote places outside the Earth, their language usage will certainly diverge. Interestingly, the isolation may slow the rate of language change, since contact with speakers of variant languages will be minimal. In these settlements, change would presumably stem mostly from situational novelty, intrinsic language characteristics, initial pool of variation among the language users, and chance elements.

References to synchronic language variation, especially sociolinguistic factors,

include [2], [5], [10], and [14].

Languages and Groups

We can often identify a person's important group memberships by the way the person speaks or writes. The basic language used will usually indicate society of origin. Saying "eh..." in a certain way means the person comes from Canada. Using certain technical slang and trailing sentences off will identify an aerospace technical professional. The notion of a "language", which is verbal, can be extended to the more general concept of a "semiotic", or a system of signs that may also be nonverbal. Manner of gesturing, for example, can reveal group allegiances.

The relation of language or semiotic systems to groups and their culture should be of concern in the design of machine intelligence systems in several ways. First, meanings may be clustered in certain ways in a given group [11], or assumptions about what is implied by certain statements may vary with the culture. Additionally, and somewhat more subtly, proprieties and niceties of presentation will probably also vary from group to group. (This point holds not just for language groups but for occupational groups, who may for example prefer different kinds of computer displays or even have feelings about the proper degree of terseness in communication.)

The conditioning of meanings by culture is by no means understood in detail, nor is the related matter of the exact effects of language on thought (and vice versa). Language and group-based differences in assumptions and implications are certainly real, though not always transparent to the participants. 'Dialectal' differences in understandings between persons who speak the same language but have different backgrounds and group-based goals, may be especially hard to detect. As a partial remedy, perhaps heuristic "expert systems" could be prepared to advise Space workers who have to deal with persons from other groups.

It should be mentioned that, although comprehensive machine translation between languages has to be at least as difficult as the analysis of a single human language, interest in this enterprise has revived over the past decade [13].

Special-Purpose Semiotics

Short of full-scale language translation by machines (a task that in its fullness may be better handled by people who spend some time interacting in another culture), we might settle for machine aids to some of the language problems that may arise in Space work.

It may be possible to devise small sets of pictorial or acoustic symbols for important matters, that can be understood with little familiarization by persons from diverse backgrounds. However, one should note that pictorial conventions to some extent differ between cultures, as apparently also do "acoustic icons" [6].

Since language processing is in fact very difficult and demanding of computer and machine storage resources even when fairly well understood, some computational linguists have aimed to characterize *sublanguages* for special content areas. Montgomery and Glover, for example [12], report on a sublanguage for describing events in space missions.

In planning for the entire machine/human communication arrangements for

orbital habitats or planetary bases, it is necessary to consider the physical arrangements as well as functional importance of matters that are communicated about. Use of sound is very natural on earth, for example, but acoustic signals must be relayed electronically when humans are in airless environments, competing with other information that must be delivered and possibly running into problems of reliability. Informally developed gestural languages are easy for humans to use, as Roger Brown has shown [3], but may be constrained by protective clothing in some situations.

1. Anttila, R.   An Introduction to Historical and Comparative Linguistics.   NY: Macmillan, 1972.
2. Bauman;, R. & Sherzer, J. (Eds.). *Explorations in the Ethnography of Speaking*. NY: Cambridge U. Press, 1974.
3. Brown, R. *Social Psychology: the Second Edition*.   NY: Free Press, 1985.
4. Cohen, P. SRI Technical Report, 1987.
5. Fishman, J. *Advances in Sociolinguistics*.   The Hauge: Mouton, 1972.
6. Hays, D. "Vocal Iconics". Auburn: SECOL, 1986.
7. Hays, D.   "Social Robotics for Exploratory Missions."   Seattle: Workshop on Planning for Autonomous Mobile Robots, 1987.
8. Hays, D.   "Trans-System Linguistics."   In Battistella, E. (Ed.)   Papers in Computation and Cognitive Science. Bloomington: Indiana Linguistics Club, 1985.
9. Hoenigswald, H.   Language Change and Linguistic Reconstruction,   Chicago: U. Chicago Press, 1960.
10. Jesperson, O. *Mankind, Nation and Individual from a Linguistic Point of View*. Bloomington: Indiana U. Press, 1964 (from Unwin, 1946).
11. Lakoff, G. *Women, Fire, and Dangerous Things: What Categories REveal About the Mind*. Chicago: U. Chicago Press, 1987.
12. Montgomery, C. & Glover, B. "A Sublanguage for Reporting and Analysis of Space Events." In Grishman, R. & Kittredge, R. (Eds.). *Analyzing Language in Restricted Domains*. Hillsdale, NJ: Erlbaum, 1986, pp. 129-162.
13. Nirenberg, S. (Ed.). *Machine Translation*. NY: Cambridge U. Press, 1987.
14. Trugdill, P. *Sociolinguistics*. NY: Penguin, 1983.