Topic #5: Apart from their role in support of MT,
what can IL representations be used for?

Using a Multi-Level Approach and Lexical Interlingual Forms in the NL Component of a Virtual Reality System

Clare R. Voss
Army Research Laboratory (ARL)
Adelphi, Maryland
voss@umiacs.umd.edu

The field of MT research lacks a consensus on what an interlingua (IL) is and how it is defined [Dorr and Voss (1993)]. MT system developers

in building their individual interlinguas have drawn on a variety of semantic formalisms and have made quite distinct assumptions concerning the overall design of the MT systems in which their formalisms are embedded [Voss (1996)].

Vanderlinden and Scott (1995) point out that, even given the variation that currently exists among individual ILs, the variation has been bounded indirectly: the current IL-based MT paradigm assumes a content invariance in sentence-by-sentence translation, in effect creating an IL "ceiling" above which variation in content selection for the IL does not occur.

Thus, any argument for using IL representations beyond the MT application—in "non-MT" environments—must be made narrowly, in terms of the MT system design where the IL was defined and the ceiling or level of representation at which the IL's content was established.

In this brief paper, I take the general position that MT researchers need to make available their IL definition, development and evaluation for re-use outside of MT. Specifically below I take the narrow position that two aspects of a "working IL" in the MT research of Dorr and Voss (1993,1996) and Voss (1996),

(i) the multi-level system design of PRINCITRAN, in which distinct representational languages are used for different types of knowledge, and

(ii) lexical interlingual forms, in which the NL semantics of English lexical items is represented,are directly relevant to "non-MT" applications. The support for this argument comes from current research developing a natural language (NL) processing system for a virtual reality (VR) environment, a "non-MT" application under construction at the Army Research Lab (ARL) [Gurney, Klipple, and Voss (1996)].

Both PRINCITRAN and the NLVR system have been developed with special attention to the same semantic domain, namely representing spatial expressions, NL sentences that describe locational relations between physical objects in 3-dimensional space (e.g., a helicopter at the airport). The difficulties that arise in MT in identifying the range of interpretations for as simple a sentence as,

"the mouse ran between the chairs"

also arise in the NLVR system, albeit typically with different objects, as in,

"drive the tank between the buildings".

Consider, for a moment, several possible meanings for these sentences. Does the mouse/tank move TO a place between the chairs/building and

stop? Or does the mouse/tank move PAST such a place on its way elsewhere? Or ABOUT in some path at such a location?

With respect to (i), the comparable multi-level design of the MT system and the NLVR system—i.e., where one can identify comparable levels of representation—makes it possible to designate where the ambiguity in the sentences above arise: namely the same places in the lexicon pre-runtime and in the parse trees at runtime.

Furthermore contributions from a discourse level of representation, not present in PRINCITRAN but in the NLVR system design (Gurney, Perlis, and Purang, 1995), are more readily assessed for integration back into MT, given the comparable system designs.

With respect to (ii), PRINCITRAN and, in due course, the NLVR system share a decompositional lexical semantics that distinguishes the semantic structure and the semantic content of its lexical entities. This also will make it possible to extend to both systems the cross-linguistic insights from research in spatial relations as well as measure phrases and aspect by Klipple (1991). It also leaves open the possibility that recent work of Asher and Sablayrolles (1995) can be tested within an NLVR system first and then, as relevant, brought to bear for translating spatial expressions in the MT system.

Ultimately the extension of IL research to "non-MT" applications ought to enable the MT community to both offer and take advantage of a wider range of software systems. Haller and Mark (1990), as just one example from the GIS (geographic information systems) community, report the significant need for an interlingua—to them, "a neutral yet expressive core of concepts"—that will support multiple representations of the same geographic object, arising both from multiple conceptualizations and lexicalizations of these objects cross-linguistically.

References

Asher, N. and P. Sablayrolles (1995) "A Typology and Discourse Semantics for Motion Verbs and Spatial PPs in French." Journal of Semantics, 12, pp. 163-209.

Dorr, B. and C. Voss (1993) "Machine Translation of Spatial Expressions: Defining the Relation between an Interlingua and a Knowledge Representation System." In Proceedings of the 12th Conference of the AAAI, pp. 374-379. Washington, D.C.

Dorr, B. and C. Voss (1996) "A Multi-Level Approach to Interlingual

MT: Defining the Interface Between Representational Languages." International Journal of Expert Systems, to appear.

Gurney, J., E. Klipple, and C. Voss (1996) "Talking about What We Think We See: Natural Language Processing for a Real-Time Virtual Environment." In Proceedings of the Second IEEE Symposium on Image, Speech, and Natural Language Systems [ISNLS], Washington, D.C.

Gurney, J., D. Perlis, and K. Purang (1995) "Active Logic and Heim's Rules for Updating Discourse Context." In Proceedings of the IJCAI Workshop on Context in Natural Language Processing. Montreal, Canada.

Haller, S. and D. Mark (1990) "Knowledge Representation for Understanding Geographical Locatives" In Proceedings of the 4th International Symposium on Spatial Data Handling, Zurich, Switzerland, pp. 465-477.

Klipple, E. (1991) "The Aspectual Nature of Thematic Relations: Locative and Temporal Phrases in English and Chinese." Ph.D. thesis, Department of Linguistics and Philosophy, M.I.T.

Vanderlinden, K. and D. Scott (1995) "Raising the Interlingual Ceiling with Multilingual Text Generation." In Proceedings of the IJCAI Workshop on Multilingual Text Generation. Montreal, Canada.

Voss, C. (1996) "Interlingua-based Machine Translation of Spatial Expressions." Ph.D. thesis, Department of Computer Science, University of Maryland, College Park, MD.