Standard Language at Ford Motor Company: A Case Study in Controlled Language Development and Deployment

Nestor Rychtyckyj
Ford Motor Company
Advanced Manufacturing & Engineering Systems
Dearborn, MI
nrychtyc@ford.com

Abstract

In 1990, Ford Body & Assembly Operations introduced Standard Language as a requirement for writing process sheet assembly instructions in North America. Standard Language was created to standardize the process sheet assembly instructions and introduce consistency across the entire manufacturing spectrum at Ford Motor Company. Standard Language is a Ford-specific, restricted subset of English that is used to describe the vehicle assembly process at Ford Motor Company. This language is used as the input to the Artificial Intelligence (AI) component of the Global Study Process Allocation System (GSPAS). Process sheets written in Standard Language are read by the AI system and used to generate work assembly instructions. Since its introduction in North America, Standard Language has been deployed to Ford's assembly plants in Europe, South America and Asia. In this paper, we will discuss our implementation of Standard Language at Ford Motor Company and show how a controlled language, such as Standard Language, can be used to provide a competitive advantage in business.

1 Introduction

Standard Language was developed as a standard format for writing process descriptions. Prior to the introduction of Standard Language, process sheets were written in free form text, which caused major problems because of ambiguity and lack of consistency. The goal of Standard Language was to develop a clear and consistent means of communicating process build instructions between various engineering functions. The use of Standard Language has eliminated almost all ambiguity in process sheet instructions and has created a standard format for writing process sheets across the corporation.

The process sheet is the standard process-planning document used to convey assembly information from the initial process planning activity to the assembly plant. A process sheet contains the detailed instructions needed to build a portion of a vehicle as well as its associated part and tooling information. A single vehicle may require thousands of process sheets to describe its entire assembly. A process engineer writes the process sheet utilizing a restricted subset of English known as SLANG (Standard LANGuage). Standard Language allows an engineer to write clear and concise assembly instructions that are unambiguous and machine-readable.

Standard Language is a type of a *Controlled Language* (Huijsen 1998). Controlled Languages were developed primarily to reduce the inherent complexity and ambiguity in natural language by simplifying the language and making it easier to read and understand. Such a restricted language also has the advantage of being much more understandable and easier to parse with a computer system. Controlled Languages are typically used for technical documentation and have been applied in a wide variety of applications, including aircraft maintenance at Boeing (Wojcik et al. 1998), aircraft design at Airbus (Spaggiari et al 2003), automotive service processing at General Motors (Godden 2000), and heavy equipment manufacturing at Caterpillar (Kamprath & Adolphson 1998).

A Controlled Language defines a set of explicit restrictions and constraints on the grammar, lexicon and style of the document being produced. The major aim of these constraints is to reduce the ambiguity, redundancy, size and complexity of the language that is being written. The lexicon or vocabulary in a Controlled Language is restricted by limiting the words or terms that can be used to those that are included in a glossary. Any additions or changes to the glossary must be approved before they can be added.

The benefits of Controlled Languages are two-fold. The documentation and text that is produced by the technical writers is consistent and unambiguous across the entire organization. All of the corporate documentation follows the same format, uses a similar sentence structure and style, and utilizes a common corporate technical glossary that is accessible to everybody. The reduction in the complexity of the syntax and lexicon usually improves the readability and comprehension of the text. Controlled Languages also provide a strong foundation for the translation of corporate documentation into other languages. The use of a controlled language simplifies the translation task both for human translators and machine translation systems.

Section 2 of this paper provides background for Standard Language and describes the motivation and development process that went into this project. In Section 3, we describe the architecture of Standard Language and demonstrate its use for vehicle assembly process planning at Ford Motor Company. A description of the Machine Translation of Standard Language is given in Section 4. The paper concludes with a summation of Standard Language at Ford and provides a glimpse into our future plans for this application.

2 History of Standard Language at Ford Motor Company

The Manufacturing Process Planning System (MPPS) was the initial attempt at Ford Body & Assembly Systems to develop an integrated system to manage process planning for vehicle assembly manufacturing for all of North America. The Direct Labor Management System (DLMS) (Rychtyckyj 1999) was developed to work with MPPS as a tool to standardize process descriptions, generate detailed plant floor assembly instructions and provide consistent and accurate estimates of the effort and time required for the actual assembly process. One of the primary reasons for the development of the Direct Labor Management System (DLMS) was the requirement to standardize process sheet writing.

The first stage in the development of DLMS was to describe a process description language that could be used by the process engineers to describe the assembly instructions at a higher generic level. The process engineer would write a sheet that was not specific to any plant, but could be applied to all the plants that were building the particular vehicle. Such a language must be unambiguous and restrictive enough to enable it to be read and understood by the DLMS system. It must also accurately describe the assembly process, but be flexible enough to allow for future growth and changes to the assembly process. Standard Language was developed to fulfill all of these requirements.

The initial version of DLMS and Standard Language was piloted at the Edison Assembly plant in 1989-1990 using a process deck for the Ford Ranger. This work focused on building and adapting a framework for Standard Language that was acceptable to the process engineering community and able to develop detailed work assembly instructions and their associated MODAPTS codes (MOdular Arrangement of Predetermined Time Standards) (International MODAPTS Society 1997). The MODAPTS system has been used at Ford for over 20 years.

A specific type of verb, known as a Standard Language verb, was selected to be the focal point of a Standard Language sentence. Each Standard Language verb (we currently have 169) was carefully defined to represent a specific action and was used as the main driver for a Standard Language element or sentence. Only one Standard Language verb is valid per element, and the resulting actions from a process instruction are based on the interaction between the verb, the object, any modifiers and the tools and parts that are used in the element.

The process of building the Standard Language lexicon and populating the DLMS knowledge base with the correct resulting MODAPTS was a group effort requiring the close cooperation of the Body & Assembly Industrial Engineers, the DLMS system developers and assorted personnel from the assembly plants. Standard Language needed to be uniquely customized to the different assembly requirements needed in Final Assembly, Body and Paint. The introduction of Standard Language into Paint proved especially challenging, due mostly to the exact dimensions that were required to describe the various surfaces and parts. We solved this problem by developing formulas that calculate the work required based on the size of the part or surface that is being used. These dimensions need to be added by the process engineer at the time the process sheet is being written.

Aside from the technical challenges experienced in the development of Standard Language, there were organizational and management issues that needed to be addressed. The development of Standard Language occurred under the management of the Body & Assembly organization, whose responsibility included the assembly plants in North America. The process engineers were used to writing free-form process sheets and the introduction of a standardized system, such as Standard Language, was met with some resistance. Some process engineers did not want to use Standard Language and went to great lengths to document all of the errors and problems with the system. This kind of feedback proved to be beneficial to us in the long run as it speeded up the debugging process. The underlying design of the DLMS knowledge base enabled us to make changes and additions to Standard Language fairly quickly to ensure that the process sheets could be fixed and then released to the assembly plants. The use of MPPS, DLMS and Standard Language was rolled out to most of the North American assembly plants through the mid-1990s. Throughout this time, Standard Language was being constantly modified based on input from Industrial Engineers doing time studies at the plants and upon the introduction of parts and tools that needed to be processed correctly.

Since that time, we have created a new application called the Global Study Process Allocation System (GSPAS) to manage all of the process planning for manufacturing across all of Ford's assembly plants around the world. Standard Language has been integrated into GSPAS and is now used throughout Ford including Jaguar and Land Rover. As the system was being deployed at different locations, we had to modify Standard Language to support the terminology and manufacturing process that were required at each of these facilities (Rychtyckyj 2004, 2006).

3 Description of Standard Language:

Standard Language is a controlled language that provides for the expression of imperative English assembly instructions at any level of detail. All of the terms in Standard Language with their pertinent attributes are stored in the DLMS knowledge base in the form of a semantic network-based taxonomy. Certain word categories in the language possess specific semantics as defined by the engineering community. Verbs in the language are associated with specific assembly instructions and are modified by significant adverbs where appropriate. For example, the phrases *inspect*, *visually inspect* and *manually verify* all have different interpretations. Information on tools and parts that are associated with each process sheet is used to provide extra detail and context.

The Standard Language sentence is written in the imperative form and must contain a verb phrase and a noun phrase that is used as the object of the verb. Any additional terms that increase the level of detail, such as adverbs, adjuncts and prepositional phrases are optional and may be included at the process writer's discretion. The primary driver of any sentence is the verb that describes the action that must be performed for this instruction. The number of Standard Language verbs is limited and each verb is defined to describe a single particular action. For example, the verbs *position* and *seat* have different meanings and cannot be used interchangeably. The object of the verb phrase is usually a noun phrase that describes a particular part of the vehicle, tool or fastener. Standard Language allows the usage of modifiers that provide additional detail for those objects. The process sheet writer may use prepositional phrases to add more detail to any sentence. Certain prepositions have specific meaning in Standard Language and will be interpreted in a predetermined manner when encountered in a sentence. For example, the preposition *using* will always signify that a tool description will follow. Figure 1 below shows how the Standard Language sentence "Feed 2 150 mm wire assemblies through hole in liftgate panel" is parsed into its constituent cases.

(S (VP (VERB FEED)) (NP (SIMPLE-NP (QUANTIFIER 2) (DIM (QUANTIFIER 150) (DIM-UNIT-1 MM)) (ADJECTIVE WIRE) (NOUN ASSEMBLY))) (S-PP (S-PREP THROUGH) (NP (SIMPLE-NP (NOUN HOLE) (N-PP (N-PREP in) (NP (SIMPLE-NP (ADJECTIVE LIFTGATE) (ADJECTIVE OUTER) (NOUN PANEL))))))))

Figure 1. Example of parsed Standard Language sentence

As mentioned previously, Standard Language process sheets are validated by the AI system for correctness. The AI validation includes the following: checking the process sheet for errors, generating the sequence of steps that a worker at the assembly plant must perform in order to accomplish this task and calculating the length of time that this task will require. The DLMS system interprets these instructions and generates a list of detailed actions that are required to implement these instructions at the assembly plant level. The detailed actions generated by DLMS are used by engineering personnel at the assembly plant to allocate the required work among the available personnel. DLMS is a powerful tool because it provides timely information about the amount of direct labor that is required to assemble each vehicle, as well as pointing out inefficiencies in the assembly process.

The parser utilizes the Augmented Transition Network (ATN) method of parsing (Charniak 1987). Any process element that is not parsed successfully will then be flagged by one of the error rules that will (hopefully) suggest to the user how to correct this element. The process engineers are also trained in the proper way to write sheets in Standard Language. Other utilities, such as an on-line help facility and written documentation are provided to assist the process engineer in the writing of correct process sheets in Standard Language. The vehicle assembly process is very dynamic; as new vehicles and assembly plants are added to the system, it requires that Standard Language also evolve. Changes to Standard Language are requested and then approved by the Industrial Engineering organization; these changes are then added into the system by updating the DLMS knowledge base. Figure 2 shows the graphical user interface that is used to maintain the DLMS knowledge base. All of the associated knowledge about Standard Language terminology, tooling information, part descriptions, labor and everything else associated with the automobile assembly process is contained in the DLMS knowledge base or taxonomy. Figure 3 shows a template that described the basic structure of a Standard Language element.

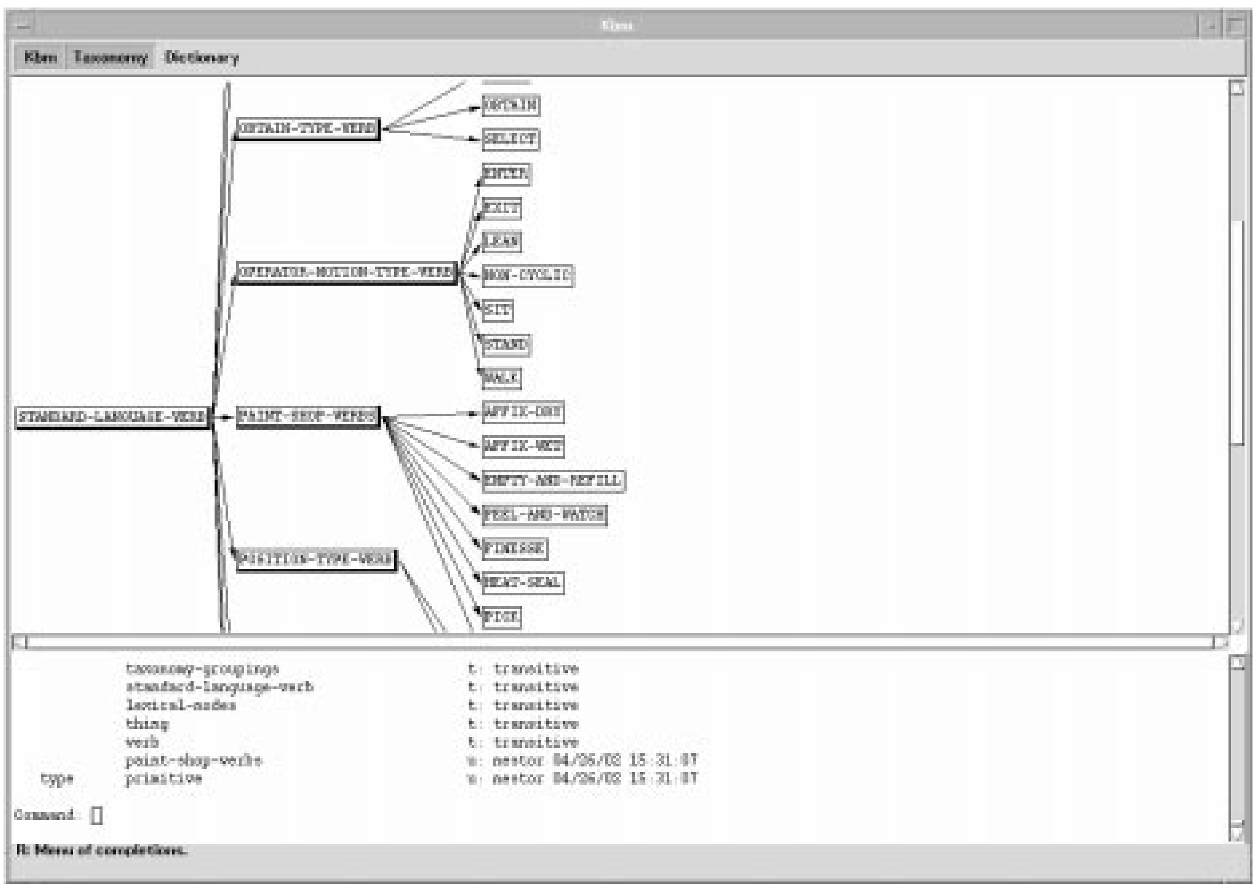

Figure 2: The Knowledge Base Manager (KBM)

Adverb	Verb	Noun Phrase	Initial Location	Intermed Location	Final Location	Fastener	Tool
	Obtain	Fuel Filler Door	From Vehicle				
Auto	Load	Sub-Assembly			To Station		
	Insert	Bolt	Into slot				Using Tool
	Verify	That Bracket is in Place					

Figure 3: Standard Language Template

4 Machine Translation of Standard Language

The restricted grammar of Standard Language and its limited vocabulary made the DLMS application a very strong candidate for machine translation. Similar applications with controlled languages have successfully utilized machine translation with a very high accuracy rate (Isabelle 1987). After an evaluation of the commercial translation packages available at the time, we decided to work with SYSTRAN (Systran 2004) for our machine translation needs. This decision was based on the wide range of languages available, the high performance of the SYSTRAN translation software, the ability to run the software under the UNIX operating system and the willingness of SYSTRAN to customize their products for our application.

The use of machine translation for Standard Language provided us with many unique challenges. The amount of data we need to translate consists of about 50,000 records for each vehicle and each language pair. Currently, we are translating over 25 such vehicle language pairs. Each sentence of text is stored as a single record in an Oracle database, which has to be retrieved, translated and then written back into the database. Issues with performance and database integrity had to be addressed, and an interface was developed by SYSTRAN and Ford to improve the translation performance.

Standard Language has adopted some grammatical structures not found in general English in order to optimize its ability to encode time and motion information. In addition, translation of standard language terms has required that similarly well-defined equivalents be designated in our target languages. Both of these have led to some specific problems in implementing machine translation.

Many of the structural decisions that were made in the development of Standard Language had a serious impact on the implementation of automated language translation technology for this application. The sentence structure in Standard Language is always imperative with the verb phrase in the beginning of the sentence. The verb phrase could also use a modifier that could impact the meaning of the verb. This modifier is in the form of an adverb, but in some cases, this word is a noun or an adjective that functions as an adverb. For example, the term “Robot Spot-weld the Object” uses the word “Robot” to modify the verb. These types of unconventional grammatical usage caused problems for SYSTRAN, which was originally developed to work with common English grammar.

A critical issue in our translation quality is in the use of part descriptions that are represented as long noun phrases: (i.e. *Side member bracket assembly medium*). If the entire phrase is not present in our technical glossary, the system will translate the phrase by splitting it up or on a word-by-word basis. In most cases, this translation is incorrect and must be manually translated by our engineers at the assembly plants.

The Ford technical terminology raised the challenge of defining translation equivalents for the well-defined terms of standard language. There are many terms that describe automotive processes and parts that are utilized only within Ford Motor Company. These terms include acronyms, abbreviations, Ford locations and other terms that cannot be translated by anybody who is not familiar with Ford. It was found that many of the terms were not understood by all of our people, as they may only be used within one department in a plant. These terms all have to be identified and translated manually so they can be added into the SYSTRAN dictionaries correctly. Problems were also caused by entries (i.e. *shotgun*) that are used informally to describe tools or equipment at the plant. Many other people may be unaware of what this term represents, and a literal translation of "shotgun" would make no sense to anybody in German or Spanish. Technical glossaries, such as those published by the Society of Automotive Engineers, are very useful in some cases, but they do not always contain a complete list of terms and can become dated and obsolete due to the rapid pace of technological progress.

Another issue with Standard Language deals with multiple spellings and misspellings of various terms. The DLMS system has a utility to allow engineers to add new terms to the system, as they are required. However, this also allows for multiple variations of the same concept due to misspellings or inconsistent usage. For example, some people would add acronyms without periods (ABS) and others would add the term with periods (A.B.S.). Over time, the knowledge base would contain quite a number of these variable spellings for the same object. Attempts were made periodically to clean up the knowledge base, but multiple terms did not become an issue until they all had to be translated.

As mentioned previously, the verbs in Standard Language are defined very concisely and unambiguously to represent a single particular action. It was not always possible to translate these verbs into other languages on a one-to-one basis to preserve their consistent meanings. The translation was accomplished only after spending considerable time on redefining their meanings in English and then translating the verb based on the most common usage in the target language. In some cases, one single English verb would have multiple translations based on its context or object that it was acting upon. Another problem arose with the use of compound verbs, which are a creation of Standard Language. A compound verb (ex: press-and-hold) was created to describe two actions with one verb that often occur together. Their usage makes it simpler for the process writers but causes complications in translations, as we are creating a new word in another language. The entire issue of defining an equivalent Standard Language lexicon for each of the target languages required considerable effort and is not entirely completed.

In Standard Language the use of articles is optional, and they are usually not written in order to save time. This leads to sentence structures, which can be easily misinterpreted by the language translation software as it is expecting complete English sentences. This problem was partially solved by modifying the parser in the AI system to add articles into the text where appropriate. Another extension to Standard Language allowed for the usage of certain adjective modifiers after the noun in order to override some attribute of the part. This structure also caused problems during translation, and the parser was modified to handle these additional problems. Abbreviations are also expanded into their full size before the translation to prevent similar errors with their meaning.

Another problem involves the use of comments within Standard Language. Any comment or remark can be included into Standard Language if it is delimited from the regular text with brackets. These comments are ignored by the AI system and do not have to conform to the Standard Language rules. The translation of free-form text and individual words is extremely unreliable and continues to cause problems. Often the comments are valid terms that should be added into the Standard Language lexicon, but the process writer often uses the commenting feature to bypass the process of adding these new terms into Standard Language. We have implemented XML tagging to assist in the translation of embedded comments. These tags work in conjunction with the Systran translation engines to pass information about the context of the terms being translated and improve the translation accuracy. Our future plans include the addition of more descriptive tags to assist in the accuracy of the translations. Figure 4 below shows an example of a Standard Language sentence that has been tagged prior to translation: The tags separate the embedded comment from the rest of the sentence and signal that these comments must be translated separately from the rest of the text. Our future plans include the addition of more descriptive tags, such as "part of speech" and context, to assist in the accuracy of the translations.

SECURE BUMPER BRACKET <comment>FOR LHS ONLY</comment> TO VEHICLE BODY USING POWER TOOL

Figure 4: Text with Embedded Tagging

We have greatly improved the accuracy of the machine translation for Standard Language. Our latest evaluation showed that 94% of the Standard Language elements are being translated correctly or acceptably into German (Rychtyckyj 2002). The evaluation of our translated text is done using the Society of Automotive Engineers (SAE) J2450 Translation Quality Metric (SAE 2002). The J2450 metric provides a consistent standard against which the quality of translation of automotive information can be objectively measured. This metric can be applied to both human and machine translation. The metric consists of seven major error categories, two sub-categories, meta-rules and numeric weights that are used to compute a translation score for a given document. Our experience with using J2450 for machine translation evaluation has been very positive in terms of getting consistent results. We have also found that most of the remaining translation errors are caused by problems in the English source text. The source text is valid Standard Language, but it is ambiguous or not understandable in regular English. We have shown that we can improve the translation accuracy by making corrections to the dictionary, modifying the Standard Language text before translation and doing a better job of translating the non-Standard Language comments. It is important to note that in some ways, despite the restricted grammar, Standard Language is more difficult to translate than regular colloquial English. This is due to the specialized terminology, the ungrammatical sentence structure and the style of the text.

5 Conclusions

In conclusion, it can be said that Standard Language has been an overwhelming success as measured by almost any metric. Previously, process sheets were written in free-form English and then sent to the assembly plants for implementation. The quality and correctness of process sheets differed greatly based upon which engineer had written a particular sheet. There was no standardization between process sheets. Industrial engineers at the assembly plants would be forced to implement work instructions based on various styles of process sheets. The introduction of Standard Language and AI forced the engineering community to write consistent process sheets and reduced the number of errors and incomplete processes that would have to be fixed at the assembly plants. It also provided a tool to investigate and estimate the labor and work required to plan for future "study" vehicles ands created a baseline of valid processes that could be adapted for other similar vehicles.

Standard Language is being used by hundreds of engineers throughout Ford and supports almost every single vehicle program that is being developed throughout the world. Standard Language has proven to be adaptable for use far outside of its original scope and has been successfully used to write thousands of process sheets. The development and maintenance of Standard Language continues to support the ever-changing and complex world of vehicle assembly process planning

The use of a controlled language in conjunction with Machine Translation has allowed us to translate our manufacturing processes for plants in Europe and South America. We feel that Standard Language will remain an integral part of Ford's manufacturing processes well into the future.

References

Charniak, E., Riesbeck, C., McDermott, D., Meehan, J., 1987. *Artificial Intelligence Programming*, Lawrence Erlbaum Associates.

Godden, Kurt (2000), "The Evolution of CASL Controlled Authoring at General Motors" in the *Proceedings of the Third International Workshop on Controlled Languages (CLAW 2000),* April 29-30, 2000 Seattle, WA, pp. 14-19.

Huijsen, Willem-Olaf (1998), "Controlled Language – An Introduction", in the *Proceedings of the Second International Workshop on Controlled Language Applications (CLAW 98)*, May 21-22, 1998, Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, pp. 1-15.

International MODAPTS Association. (1997), *Modapts Modular Arrangement of Predetermined Time Standards- The Language of Work.*

Isabelle, P.: “Machine Translation at the TAUM Group,” in M. King (ed.), Machine Translation Today, Edinburgh University Press (1987) 247-277

Kamprath, C., Adolphson, E., (1998), "Controlled Language for Multilingual Document Production: Experience with Caterpillar Technical English" in the *Proceedings of the Second International Workshop on Controlled Language Applications (CLAW 98),* May 21-22, 1998, Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, pp. 51-61.

Rychtyckyj, N., (1999), “DLMS: Ten Years of AI for Vehicle Assembly Process Planning”, *AAAI-99/IAAI-99 Proceedings*, Orlando, FL, July 18-22, 1999, pp. 821-828, AAAI Press.

Rychtyckyj, N., (2002), "An Assessment of Machine Translation for Vehicle Assembly Process Planning at Ford Motor Company", *Machine Translation: From Research to Real Users, 5th Conference of the Association for Machine Translation in the Americas, AMTA 2002 Tiburon, CA, USA, October 2002 Proceedings*, pp. 207-215, Springer-Verlag.

Rychtyckyj, Nestor (2004), "Maintenance Issues for Machine Translation Systems", *Machine Translation: From Real Users to Research: Proceedings of the 6th Conference for the Association of Machine Translation in the Americas,* September 28- October 2, 2004, Washington DC, Springer-Verlag, Lecture Notes in Computer Science, vol. 3265, pg. 252-261.

Rychtyckyj, Nestor, (2006), "Machine Translation for Manufacturing: A Case Study at Ford Motor Company" to be published in the Proceedings of the 18th Innovative Applications of Artificial Intelligence Conference (IAAI-2006), Boston, MA, July 18-20, 2006, (forthcoming).

Society of Automotive Engineers (SAE), (2002), *J2450 Translation Quality Metric.*

Spaggiari, L., Beaujard, F., Cannesson, E., (2003), "A Controlled Language at Airbus", *Proceedings of the Joint Conference combining the 8th International Workshop of the European Association for Machine Translation and the 4th Controlled Language Application Workshop*, pp. 151-159.

SYSTRAN Language Translation Technologies, (2004), www.systransoft.com , Paris, France.

Wojcik, R., Holmback, H., Hoard, J., "Being Technical English: An Extension of AECMA SE Beyond the Aircraft Maintenance", in *Proceedings of the Second International Workshop on Controlled Language Applications (CLAW 98),* May 21-22, 1998, Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, pp. 114-123.