[From: *CLAW '96*, Universiteit Leuven]

The Controlled Automotive Service Language (CASL) Project

Linda Means and Kurt Godden
General Motors Research and Development Center
Warren, Michigan, U.S.A.
rneans@gmr.com, godden@gmr.com

ABSTRACT

The General Motors Controlled Automotive Service Language (CASL) project will provide a controlled language designed for translatability, software tools for authoring service information, and English-to-French translation with the Metal machine translation system. The resulting authoring and translation system will greatly reduce cost and time of translation, and improve the consistency and readability of the English source documents. This paper discusses technical issues as well as the organizational challenges of centralizing and standardizing the authoring process at General Motors.

INTRODUCTION

GM publishes approximately 1.5 million words of service information per year in English for North American products (29 service manuals with 3,500 pages per book, plus technical service bulletins). GM service manual text tends to be verbose, complex, and ambiguous, resulting in a high cost to translate. The complexity also impedes usability of the domestic manuals by English-reading technicians.

Recently, General Motors North American Operations has developed an Electronic Service Information (ESI) system to centralize the authoring, data management, and electronic delivery of information that supports the servicing of GM NAO products in the U.S. and abroad. During the development of ESI, GM studied issues pertaining to readability and translatability of service information, hoping to take advantage of the centralized authoring to improve data quality and translation efficiency. The Controlled Automotive Service Language (CASL) Project was the outcome of this investigation.

At General Motors, service information for different products or subsystems traditionally has been created by numerous organizations, with little sharing of processes, systems, and data. Although multiple organizations will continue to author service information within ESI, the implementation of a centralized ESI system enables standardization and sharing of the data, which was not possible in the disjoint environment of the past. In the past, the translation of service information has been also decentralized and outsourced to a variety of translation agencies, which similarly deters efficiency and consistency of translation.

The CASL project (pronounced like 'castle') strives to develop and implement technology, processes, and an infrastructure to produce source documents which are easily readable and translatable. Initial planning of the CASL project considered various strategies for maximizing translation efficiency, and identified their compatibility with current and potential organizational processes within the GM service authoring community. The resulting CASL system will be positioned to provide technology and processes which can be implemented effectively within the current organization, with the addition of author and translator training and technical support. CASL is also designed for portability to additional authoring applications in other domains.

WHAT IS CASL?

CASL Deliverables

The CASL system consists of three primary deliverables:

1. CASL English: a controlled sublanguage. The CASL project will define and implement a controlled sublanguage for service information authors. All ESI authors will be required to conform to the CASL lexical and grammatical restrictions.

CASL requires a grammar which minimizes ambiguity, enhances readability and translatability, yet is expressive enough to satisfy authors' needs. It is based on the following factors:

- General principles of readability and translatability:
  - Readability is a factor of the complexity of sentence structures, the amount of ambiguity, and the use of standardized vocabulary.
  - Translatability is a factor of the readability level, and the number of words.
- Grammatical constraints which are optimized for the particular capabilities and limitations of the Metal machine translation (MT) system.
- Formal standards, both internal to GM and within the automotive industry.
- Current usage in GM service manuals.

A preliminary version of the CASL rule set contains 95 restrictions, covering lexical, syntactic, and punctuation phenomena.

2. CASL Conformance Checker. CASL will develop a software tool which checks documents for conformance to CASL restrictions. This conformance checker will be constructed by adapting the existing Metal parser and grammar rules. The conformance checker will be integrated into the ESI authoring environment (a customized version of the ArborText AdeptEditor for SGML). Conformance checker processing will be performed on one or more servers which communicate with authors' PCs via network or modem.

Input to the conformance checker consists of an ESI Service Information Object (SIO), containing up to several paragraphs of text. The conformance checker reformats input as needed, performs a complete parse, and marks up sentences for user review. Conformance checker output is the SIO with markups for CASL violations. The user interface will allow authors to review and edit conformance checker output sentence by sentence, and to resubmit edited sentences to the conformance checker for review.

The conformance checker will be the gatekeeper for all text going into the ESI database. The ESI database will require CASL conformance for all data objects which are released into the database. The CASL conformance checker will take advantage of Metal's broad-coverage grammar to enable analysis of sentences which fall outside the range of the CASL grammar, and to recommend revisions to non-conformant text whenever feasible.

User feedback will vary, depending on the result of the analysis and the type of error:

- Correction. CASL will recommend changes, which must be confirmed by the user.
- Diagnosis. CASL will indicate violations of grammatical restrictions, and highlight the text in which they occur.
- Global feedback. CASL will produce a summary of the occurrence of various phenomena in an entire document. This will call the user's attention to questionable phenomena, such as passive voice, where we do not want to require the user to rewrite each occurrence, but we do want to alert the user to possible overuse of the construction. These summary statistics will also be used by the CASL team to help drive quality control improvements to the system as well as to suggest areas where increased author training is needed.

The conformance checker will have the following capabilities:

Lexical:

- Recognize words which fall outside of the approved CASL vocabulary.
- Recommend approved CASL synonyms for some non-CASL-approved words. For example, use 'generator' instead of 'alternator'.
- Provide usage examples to show right and wrong usage. For example, if 'slip' is allowed only as a verb, and the user submits a sentence with the noun phrase 'belt slip', display right and wrong usage examples for 'slip', and recommend the synonym 'slippage'.

Syntactic:

- Identify syntactic structures which violate CASL restrictions. For example, CASL generally prohibits gapped constructions.
- Inform users of syntactic errors.
- For specific error types, provide right and wrong usage examples.
- For errors with a strong diagnosis, provide a recommended rewrite for the sentence. For example, CASL prohibits the use of the present progressive aspect. The conformance checker may be able to automatically correct these occurrences to simple present tense.

3. CASL English-French translation. Translation suppliers will operate Metal as a translation tool. Newly authored text, for which no previous translation is found in the translation memory, will be translated by Metal. Postedited translations will be stored in the ESI database for future use in the translation memory.

CASL Scope

The initial scope of CASL includes authoring of all GM North American service manuals and service bulletins in CASL English, and English-to-French translation with

Metal. More language pairs may be added after English-French. Approximately 400 authors at dozens of sites author GM service information, which includes 29 service manuals (published annually), service bulletins (published as needed), and training materials for service technicians.

CASL Benefits

The most obvious benefit of the CASL system will be the cost savings for English-to-French service manual translations. However, other benefits which we identified but could not measure argue strongly for the implementation of the CASL system. In addition to cost savings, we have high confidence in predicting the following additional benefits for GM in the U.S. as well as worldwide:

- Improved quality of source documents.
- Increased usage of service manuals by technicians.
- Improved ability to service vehicles correctly the first time.
- Lower warranty costs.
- Higher customer satisfaction.
- Increased sales.

Pertaining specifically to translation of service manuals, the following additional benefits are expected:

- Reduced lead time in producing translations.
- Compliance with language laws in Quebec and other countries.

CASL INTEGRATION WITH ESI

CASL will be integrated with the ESI system and process in various ways:

Pre-CASL Training

Because the newly-implemented ESI authoring system requires extensive initial training for new users, the current training program provides an opportunity to introduce authors to the concept of controlled English for readability and translatability. Two days of the 7-day ESI training course are currently devoted to instruction in controlled English. This pre-CASL language training should result in reduced complexity and ambiguity in the text which is created before the implementation of the CASL authoring system. It also will help authors to become accustomed to new CASL standards which deviate from traditional practice. For example, GM service information authors have traditionally omitted many articles before concrete nouns in service manuals, and CASL will require the use of these articles. Although this is, in principle, a simple restriction to learn, the old telegraphic style is so habitual that it may take authors some time to recalibrate their writing style.

The pre-CASL training course prescribes a writing style which is clear, concise, and direct. The guidelines which were developed for this pre-CASL phase cover the topics which are most critical to translatability and readability, and which are most easily understood, remembered, and applied by authors with no supporting software tools. An

example would be a restriction on passive verb constructions, and exercises to help authors identify and rewrite passive sentences. The instructor presents rules covering various grammatical issues, and includes practice sessions in which class participants apply the controlled style in rewrite exercises from existing GM service manuals.

Database Enhancement

The ESI database was not originally designed to track foreign-language versions of data objects. To prepare for eventual implementation of translation memory and electronic data exchange with translation suppliers, the database is currently being enhanced to accommodate multiple languages, and to facilitate electronic communications with translators.

Production Authoring and Machine Translation Systems

CASL will be implemented in two stages. First, the conformance checker will be phased in among authors. This implementation will be supported by extensive author training in CASL English, and in the use of the conformance checker software. Then, the machine translation system will be provided, with training, to translation suppliers.

Quality Control

A quality control system will be developed to assess the performance of the conformance checker and the MT system, and drive changes to both systems. This system will use statistical process control charts to monitor errors in authoring, parsing, and translation. Feedback from the charts will help system developers to improve author training, and to identify needed enhancements to the Metal grammar and parser.

ORGANIZATIONAL SYNERGIES

To be successful, a CASL-like system requires an appropriate infrastructure, in which the work of authors and translators is closely managed. At GM, CASL will impose new restrictions on the work of authors and translators, and will drive some changes in their work patterns.

To the greatest extent possible, we attempt to make CASL fit into the existing process, and avoid changing existing processes to accommodate the new CASL system. This is particularly important in our environment, where the authoring process encompasses a large number of organizations, and it can be difficult and risky to manage process changes.

The Authoring Process

CASL's largest liability is probably the makeup and organization of the authoring community for GM service information. However, the CASL project tries to recognize weaknesses in the system and offset them with CASL processes and technology.
ESI's 400 authors work for numerous organizations at dozens of sites. They are generally not trained as technical writers; many were hired because of their experience as automotive technicians, which provides the technical expertise needed for the authoring task. Because of their inexperience as technical writers, CASL authors will need some basic grammatical training in order to understand and apply the restrictions imposed by CASL English. Because of their geographical and organizational dispersion, training and continued support for authors is a challenge.

The authors are new to ESI as well, and have a large task in learning how to use the database and the SGML editor. CASL imposes another large learning task, in writing CASL English, and in use of the conformance checker. ESI provides a centralized data management system, without which CASL would be impossible. CASL, however, will impose an even greater degree of standardization on the ESI authors.

CASL uses various strategies to provide a system which is compatible with the existing authors and authoring processes:

- Provide a conformance checker with sufficient diagnostic, explanation, and automatic correction capabilities to assist authors who are not very knowledgeable in English grammar.

- Develop an author training course which relies more heavily on learning by example than on grammatical knowledge.

- Coordinate CASL training with other training activities for service information authors.

Retention of Trained Authors

Because of the extensive training of ESI/CASL authors, as well as the further on-the-job training which will be gained through experience with the system, it is important to recognize the retention of trained authors as a management priority.

Editorial Support

Although the CASL conformance checker can be viewed as an electronic editor, a human editing process is also needed to supplement CASL conformance checking. Human editors are needed to verify that sentences are readable and CASL-compliant, and to authorize exceptions to CASL restrictions when needed. This editorial process will assure application of the CASL standard, which assures the readability and translatability of CASL text. Any deviation from CASL will degrade both readability and translatability.

Production Schedules for Authoring and Translation

We anticipate that the introduction of the CASL conformance checker will lengthen the time required for authoring, especially when the system is newly implemented, and to a lesser extent in the long term. Authors currently work under demanding production schedules, and would find it difficult to meet existing deadlines with the addition of the CASL authoring tool. The introduction of a human editorial process will further lengthen authoring time. For these reasons, CASL cannot succeed well without an extended authoring schedule.

On the other hand, the shift from human translation to machine translation results in a considerable speed up in the translation process. This, taken with the lengthier authoring process, indicates the need for organizational acceptance of a shift in the total production schedule for service manuals, across composition and translation. This need not represent a longer total schedule; it simply allocates longer authoring time prior to translation, but reduces time required for translation at the end of the production cycle.

This shift of time from translation to authoring may have significant advantages for English source documents as well as translations:

- The ESI authors will have more time to perfect the original English documents, improving the overall quality of source and translated documents.

- The later onset of translation allows late engineering changes to be reflected in the final service manual. The additional time will also allow for the documents to be more thoroughly validated before publication.

- This closer match between the vehicle and the manual will result in the need for fewer technical bulletins (which address issues that are not covered in the service manuals). The increased integrity of the service manuals will improve technician performance at dealerships, resulting in lower warranty costs and higher customer satisfaction, ultimately leading to higher sales.

Extralinguistic Issues

CASL English is designed to address the most critical and easily analyzable factors which affect translatability at and below the sentence level. However, there are various extralinguistic issues which impact readability and translatability as well, which cannot be managed effectively through natural language software.

GM service manuals tend to be quite verbose. The information content has not traditionally been prescribed in any standard fashion. Because technicians in GM dealerships have varying levels of technical expertise, service manuals have grown to include detailed information on theory of operation to accommodate technicians with less experience and training. Nomenclature, writing style, and level of detail varies from one manual to another, and even within a single manual, because past authoring processes have not coordinated consistency in these areas.

CASL cannot control these factors in the authoring process. However, another GM initiative is developing a training workshop for simplified technical authoring, which will address information content, organization, flow, and consistency. It will also encourage a shift from a text-intensive manual to a graphics-intensive manual. These strategies for limiting the quantity of text will also improve translation efficiency and readability. The training workshops for this program will bring authors into a classroom to work on their own materials. Simplified writing work done in class will be based on CASL guidelines, and will result in CASL-conformant models for the authors' future reference.

One of the cornerstones of the ESI philosophy is the extensive use of "plagiarism" — reuse of previously authored material with little or no modification where possible. Reusing text across service manuals or across model years can improve author efficiency, and combined with the use of a translation memory tool, can reduce the amount of source material requiring translation as well. An additional benefit of extensive "plagiarism" is the side effect of greater consistency in style and nomenclature.

A Single, Flexible Process

The CASL system offers the flexibility needed to author, manage, and translate all service information with a single process. Because of the organizational complexity combined with demanding production schedules for authoring, we anticipate situations in which newly authored text cannot be made CASL-conformant, and non-CASL text will be released into the database. For this reason, a broad-coverage MT system such as Metal is advantageous. Although non-CASL-conformant text will result in lower MT accuracy, we

can still use the identical process of authoring, data management, and translation for CASL and non-CASL text.

Similarly, the identical translation process can be used for language pairs for which we have MT (initially English-French), and for non-MT language pairs. All language pairs can utilize Metal's translation memory capability, and the same work flow and editing tools can support MT postediting as well as non-MT human translation. Translators operating a CASL workstation will take ESI data as input, invoke translation memory, and the translator will process the remaining text either by postediting a Metal translation, or by translating within the Metal postediting environment.

Translation Outsourcing

CASL considerations call for a reformulation of our relationship with translation suppliers. GM currently outsources all translation of service information to numerous outside suppliers. Data transfer between GM and its translation suppliers is usually in the paper medium (not electronic). Content which has not changed since the previous year is not identified by GM. The translator can achieve some efficiency by manually identifying unchanged text and reusing the previous translation.

With the advent of ESI and CASL, the relationship with translation suppliers will change significantly. Translators will operate the Metal system to invoke translation memory, and to translate and postedit text in supported MT language pairs. The combination of translation memory and high-quality MT will greatly reduce the actual amount of translation labor required.

Because CASL MT translations will be produced at a much lower cost than full human translation, GM will expect substantial price reductions from translation suppliers. But while the per-unit price for translation will be lower, the increased capacity for translation will enable us to offer more volume to our translators to offset the reduced per-unit price.

With CASL translation capabilities, we will ask translators to use our technology, lower the unit price for translation, and comply with possible restrictions on the use of the technology. At the same time, it is important that we maintain our independence from suppliers to enable agility in awarding new contracts. This new set of requirements represents a significant change in our interaction with translators, and will call for scrutiny in planning a mutually-agreeable new supplier relationship.

CONCLUSIONS

CASL represents a progressive business strategy toward GM's globalization efforts. The pre-CASL process allocates translation jobs with little control over price, quality, consistency, and duplication of effort. CASL will centralize and commonize the translation process, and introduce quality controls, at a greatly reduced cost. CASL has the added benefit of providing simplification in domestic service manuals used by over 95,000 English-reading technicians in the U.S.

Because a system like CASL impacts the organization as well as its technologies, project plans must address not only the technology, but also the entire organizational vision needed to support the technology. The decision to adopt a CASL-like system must incorporate a commitment to making the organizational adjustments needed to ensure its success.

ACKNOWLEDGMENTS

The authors would like to thank Metal developers Uus Knops, Gregor Thurmair, and Johannes Ritzke for their kind and patient cooperation in CASL planning and development. CASL could not be developed at General Motors without the support and efforts of the GM Service Technology Group, GM Canada Ltd., and Electronic Data Systems. We also extend our appreciation to Lutz Graunitz of Max Delta Corporation for his unfailing service to CASL.