Lecture Notes in Artificial Intelligence 2308, Springer, 2002.

Symbolic Authoring for Multilingual Natural Language Generation

Ion Androutsopoulos, Dimitris Spiliotopoulos, Konstantinos Stamatakis, Aggeliki Dimitromanolaki, Vangelis Karkaletsis and Constantine D. Spyropoulos

Software and Knowledge Engineering Laboratory
Institute of Informatics and Telecommunications
National Centre for Scientific Research (NCSR) "Demokritos"
P.O. Box 60228, GR-153 10 Ag. Paraskevi, Athens, Greece
e-mail: {ionandr, dspiliot, kstam, adimit, vangelis, costass}@iit.demokritos.gr

Abstract. We describe the symbolic authoring facilities of the M-PIRO project. M-PIRO is developing technology that allows personalized multilingual object descriptions, in both textual and spoken form, to be produced from symbolic information in a database and small fragments of text. The technology is being tested in the context of electronic museums, where a prototype that produces dynamically multilingual exhibit descriptions for presentations over the web has already been developed. This paper focuses on M-PIRO's authoring subsystem, which allows domain experts with no language technology expertise to configure the system for new applications. The authoring facilities allow the experts to define or modify the structure of the underlying database, its contents, and the system's domain-dependent linguistic resources. Previews of the generated texts can also be produced during the authoring process to monitor the content and quality of the resulting descriptions.

1 Introduction

This paper presents the symbolic authoring facilities that are being developed within the M-PIRO project.[1] Drawing upon techniques from natural language generation [17], speech synthesis, and user modeling, M-PIRO is developing technology that allows personalized descriptions of objects to be generated dynamically in several languages, in both textual and spoken form, starting from symbolic, language-independent information in a database, and small fragments of text. The resulting technology is expected to have a wide range of applications, from electronic sales catalogues to computer games. During the project, it is being tested in the context of

[1] ~~M-PIRO (Multilingual Personalized~~ Information Objects) is a project of the Information Societies Programme of the European Union, running from February 2000 to January 2003. The project's consortium consists of the University of Edinburgh (UK, co-ordinator), ITC-irst (Italy), NCSR "Demokritos" (Greece), the University of Athens (Greece), the Foundation of the Hellenic World (Greece), and System Simulation Ltd (UK).

electronic museums, to enhance web-based interaction with exhibit collections and speech-enabled tours in virtual reality.

Although the project is still in progress, large-scale prototypes have already been implemented, and they will be used in this paper to highlight the functionality of the emerging technology. Figure 1 shows an example from M-PIRO's current web-based prototype. Visitors select exhibits from a catalogue that contains thumbnail images, and the system replies with dynamically generated descriptions of the exhibits. Apart from the sentence that describes the wedding scene, all of the text in Figure 1 has been generated automatically from non-linguistic information in the database. The descriptions can also be generated in Italian and Greek, as demonstrated in Figure 2, from the same underlying database, reducing dramatically translation costs. Furthermore, the descriptions are customized according to what the visitor has already seen, avoiding repeating information that has already been conveyed, and comparing, when possible, the current exhibit to previous ones. The text of Figure 1, for example, points out that the exhibit belongs to the same period as the previous one. The description is also tailored according to the user type. The prototype distinguishes between children, non-expert adults, and experts. Descriptions for children are typically shorter, while expert descriptions contain, for example, additional references to related articles, and avoid explaining common archaeological terms.

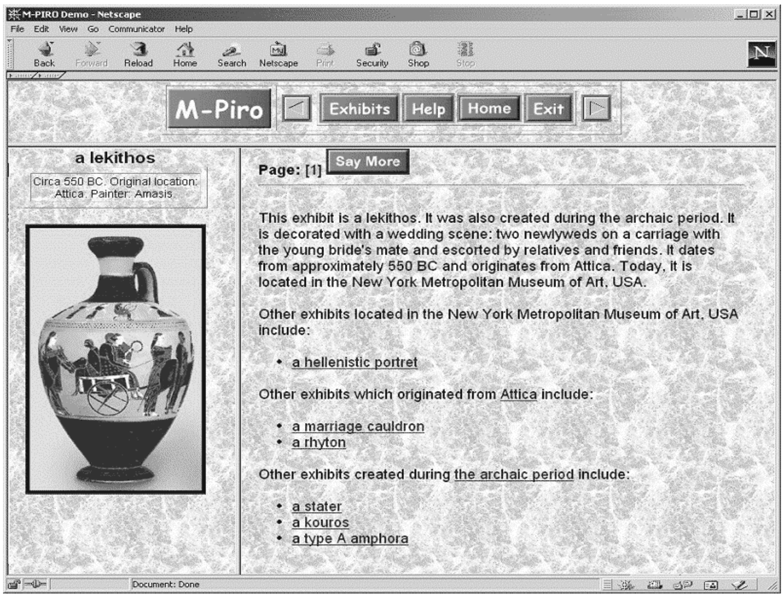

Fig. 1. A dynamically generated exhibit description in English

M-PIRO builds upon the ILEX natural language generation system [12, 13], which was originally used to produce dynamically exhibit descriptions for a web-based electronic gallery of 20th century jewellery. It extends ILEX's technology by incorporating improved multilingual capabilities [5], a more modular core generation engine, and high-quality speech output [3, 19]. The latter is needed in virtual reality tours, where work has just commenced to use M-PIRO's technology as an intelligent guide. One of M-PIRO's most ambitious goals, which also distinguishes it from ILEX and other similar generation systems [4, 10], is that domain experts, such as curators in the context of museums, will be able to configure M-PIRO's technology for new application domains, e.g., new museum collections or collections outside the museum context, without the intervention of language technology experts. This is achieved via M-PIRO's authoring subsystem, which is the focus of this paper; a broader overview of M-PIRO can be found elsewhere [1]. Although familiarity with computers and some training on the use of the authoring subsystem is still required, M-PIRO's authoring facilities constitute a significant advance compared to most natural language generation systems, where porting the system to a new domain requires programming and expertise in natural language generation.

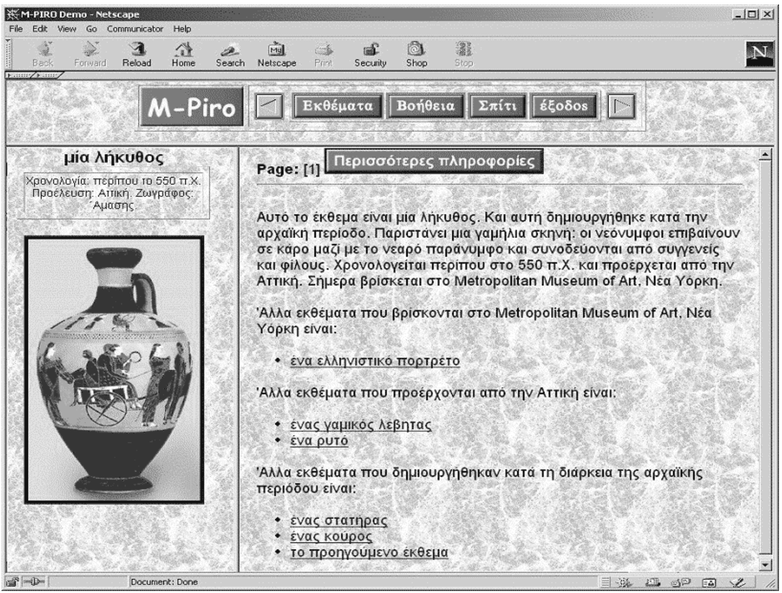

Fig. 2. A dynamically generated exhibit description in Greek

Unlike systems like KPML [2], M-PIRO's authoring subsystem is not intended to assist language technology experts in creating and maintaining domain-independent linguistic resources, such as large-scale grammars. In that sense, M-PIRO's authoring is closer to the symbolic authoring facilities of DRAFTER [14] and GIST [15]. Unlike those systems, however, M-PIRO does not target a specific application domain, and allows the domain experts, hereafter called *authors* to manipulate not only the contents of the database, but also its structure and the domain-dependent linguistic resources that control how the information of the database is rendered in natural language. This allows the authors to control, for example, the vocabulary and form of the generated sentences, as well as, in ongoing work, the rhetorical structure of the resulting descriptions [8].

Section 2 below provides more information about the role of the authoring subsystem in M-PIRO's architecture. Sections 3, 4 and 5 then discuss in more detail some of the facilities that the authoring subsystem provides, namely facilities that allow the authors to manipulate the underlying database, domain-specific aspects of sentence planning, and the domain-dependent lexicon, respectively. Section 6 concludes with targets for future work, which include evaluation plans and ways to re-use information in existing databases.

2 System Architecture and the Role of the Authoring Subsystem

Figure 3 illustrates the role of the authoring subsystem in M-PIRO's architecture. Once the user has selected an object, the system retrieves from the database all the relevant information, and produces an appropriate textual description of the object using natural language generation techniques, to be discussed briefly below. In virtual reality environments, the description is then passed to a speech synthesizer, which produces the audio output, exploiting additional markup made available by the generation components, much as in [18].

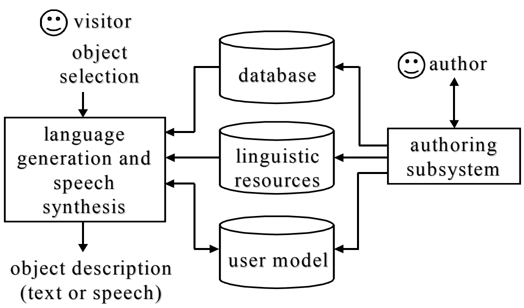

Fig. 3. The authoring subsystem in M-PIRO's architecture

Many of the linguistic resources on which the generation process relies, most notably its systemic grammars [5, 6], are to a large extent domain-independent. Some of these resources, however, are domain-specific, and one of the roles of the authoring subsystem is to allow authors to modify them for new application domains, hiding many of the underlying linguistic complexities. A second role of the subsystem is to allow the authors to manipulate the structure and contents of the database, establishing links between database constructs and linguistic resources where necessary. The third role of the subsystem is to help the authors define the types of visitors and their properties. Among other things, this includes defining stereotypes that indicate the educational value and interest of the various facts in the database for each visitor type. M-PIRO's user modeling mechanisms are based on those of ILEX, which are described in [12] and [13]. We will highlight some of the user modeling tasks that the authoring subsystem is faced with, but since work on these aspects of authoring is still in progress, this paper will focus on facilities that manipulate the database and domain-dependent linguistic resources.

To obtain a clearer view of the authoring tasks, let us now examine briefly the stages of the generation process in M-PIRO, as outlined in Figure 4; we ignore in the rest of this paper issues related to speech synthesis. The input to the generation process is the database, as shaped by the authors using the facilities that allow them to manipulate its structure and contents, and the object to be described. The first stage of the generation process, called *content selection*, is concerned with the selection from the database of the most appropriate facts to be conveyed to the visitor. It exploits user modeling information, such as the stereotypes mentioned above and the interaction history of the visitor, which shows the facts that have already been conveyed. The next stage, *document planning*, outputs an overall document structure, which specifies, for example, the desired sequence of the facts in the generated description, and their rhetorical relations; for example, whether a fact amplifies or contrasts another one [7, 8].

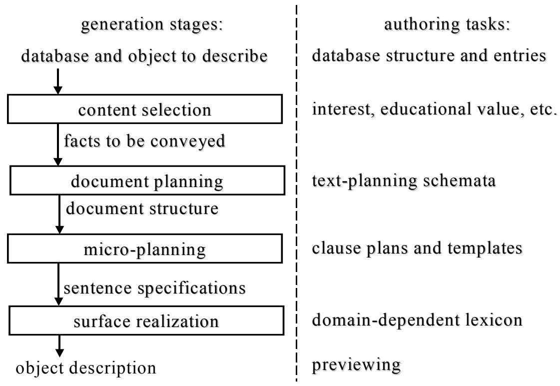

Fig. 4. Generation stages in M-PIRO and the corresponding authoring tasks

M-PIRO has inherited from ILEX a variety of domain-independent document planners, which are being extended to allow the authors to specify domain-dependent schema-like planning rules [9] to capture structural characteristics of object descriptions in particular domains. Descriptions of museum exhibits, for example, typically start with information about the type and creation period of the exhibit. The curator of a collection of coins may wish to specify that descriptions should then proceed with a description of what the two sides of the coin depict, followed by information about the material and style. Work on this aspect of authoring is just starting in M-PIRO, and will not be discussed any further. In contrast, the authoring facilities that are associated with the next two stages of the generation process, *micro-planning* and *surface realization*, are more fully developed.

Micro-planning specifies in abstract terms how a fact can be expressed as a clause in each language; for example, which verb to use, in what tense, and which argument of the fact should be rendered as subject or object. The authoring subsystem allows this information to be specified in two alternative forms, *clause plans* and *templates*, to be discussed in Section 4. Micro-planning also includes the generation of referring expressions, also to be discussed in Section 4, and processing that determines which facts can be aggregated in a single sentence. M-PIRO employs the aggregation rules of ILEX (see [13]), which are domain-independent, and hence require no input from the authors.

The last stage, surface realization, is responsible for producing the final textual form of the descriptions. This includes producing the appropriate word forms (e.g., verb tenses) based on the sentence specifications output by micro-planning, placing the various constituents (e.g., subject, verb, object, adverbials) in the correct order, accounting for number and gender agreement, etc. Surface realization is based on large-scale systemic grammars [6], one for each supported language, that were constructed using ILEX's English grammar as a starting point [5]. While the grammars are domain-independent, a part of the lexicon that they employ, called the *domain-dependent lexicon*, needs to be tuned when the system is ported to a new domain; related authoring facilities will be discussed in Section 5.

Finally, it is important to be able to preview the resulting object descriptions, to monitor the content and quality of the generated texts. The authoring subsystem allows previews to be generated during the authoring process, a point that will be illustrated in following sections. In effect, this introduces a form of interactive symbolic authoring, whereby changes in the symbolic description of the domain and the linguistic resources are immediately reflected on the generated object descriptions.

3 Database Structure and Entries

Let us now examine the facilities that are available to manipulate the structure and content of the database. An entity-relationship model is assumed; i.e., the database is taken to hold information about entities (e.g., statues, artists) and relationships between entities (e.g., the artist of each statue). Entities can be concrete or abstract objects (e.g., historical periods or styles), and they are organized in a hierarchy of entity types, as illustrated in Figure 5. In this example, *exhibit* and *historical-period* are basic entity types; the *exhibit* type is further subdivided into *vessel*, *statue*, and *coin*. Each entity belongs to a particular entity type; for example, *exhibit 2* is a *kouros* and, therefore, also a *statue* and an *exhibit*. To make the authoring subsystem easier to use, we have opted for a single-inheritance hierarchy, although the underlying generation engine can also handle multiple inheritance. There are also mechanisms to link the basic entity types to the Upper Model [2], a built-in domain-independent hierarchy that contains the most common types; this allows making some aspects of the generation process insensitive to the domain-dependent hierarchy.

Relationships are expressed using fields. At each entity type, it is possible to introduce new fields, which then become available to all the entities of the type and its subtypes. For example, the *statue* type in Figure 5 introduces the field *sculpted-by*; consequently, all the entities of this type, including entities of type *kouros* and *portrait*, carry this field. The *creation-period* field is inherited from the *exhibit* type, and is, therefore, also available with non-statue exhibits; inherited fields are shown in different colour. The fillers of each field must be entities of a particular type. In Figure 5, the fillers of *creation-period* must belong to the type *historical-period*; this licenses entities like *archaic-period* and *classical-period* to be used as values of the field. The *Set?* option in Figure 5 allows a field to be filled by multiple fillers of the specified type; in the *made-of* field, this allows entering more than one material.

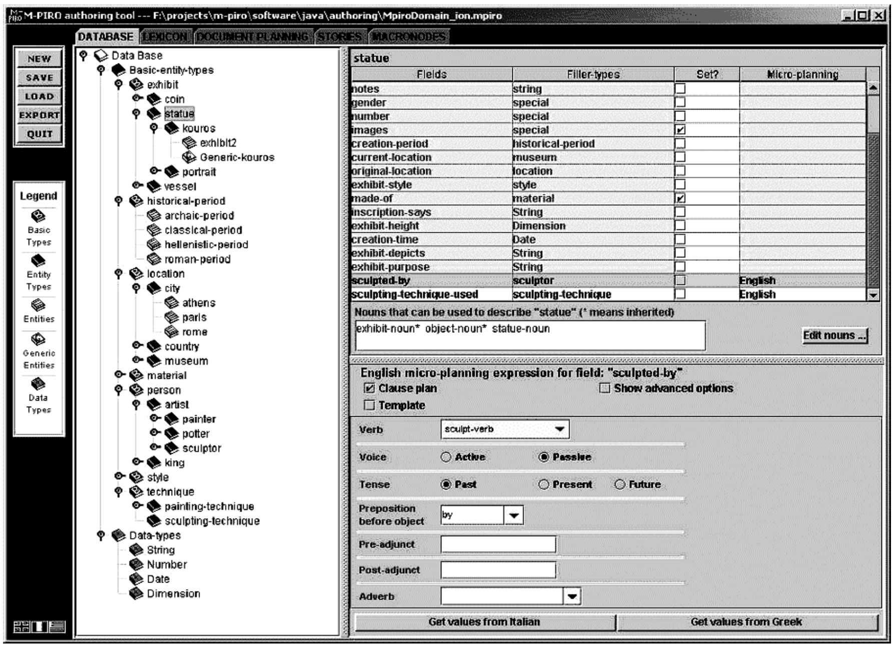

Fig. 5. The structure of the database and a clause-plan

Fields are also used to express attributes of entities, for example, their names or dimensions. Several built-in data-types are available, like *string* and *date*, and they are used to specify the allowed values of attribute-denoting fields. In Figure 5, *exhibit-depicts* and *exhibit-purpose* are string-valued. They are intended to hold canned sentences describing what the exhibits depict and their purposes; the sentence that describes the wedding scene in Figure 1 is the value of an *exhibit-depicts* field. String-valued attributes are used with information that is too difficult to express using full text generation; the drawback is that their values must be entered in all of the supported languages. Notice, however, that some of the benefits of natural language generation are still available with string-valued attributes; for example, they are assigned interest and educational values, like all the other facts in the database, and the text planning schemata can be instructed to place their string values to appropriate positions. Larger, paragraph-long canned texts can be associated with particular entities or entity types via the *stories* tab of Figure 5.

Once the hierarchy and the fields of the entity types have been created, it is possible to insert database entries about particular entities, as illustrated in Figure 6. Pull-down menus and forms guide the authors to select among the allowed values of the fields. Provided that appropriate lexicon entries and micro-plans –to be discussed in following sections– have been entered, previews of the resulting object descriptions can be generated, as shown in Figure 6.[2] The values of language-dependent fields, such as the string-valued *exhibit-purpose*, which gives rise to the sentence about Kroissos, are entered by clicking on the flags in the upper right part of Figure 6. In this example, content selection has chosen not to convey the information about the exhibit's material, because its interest and educational values are low.

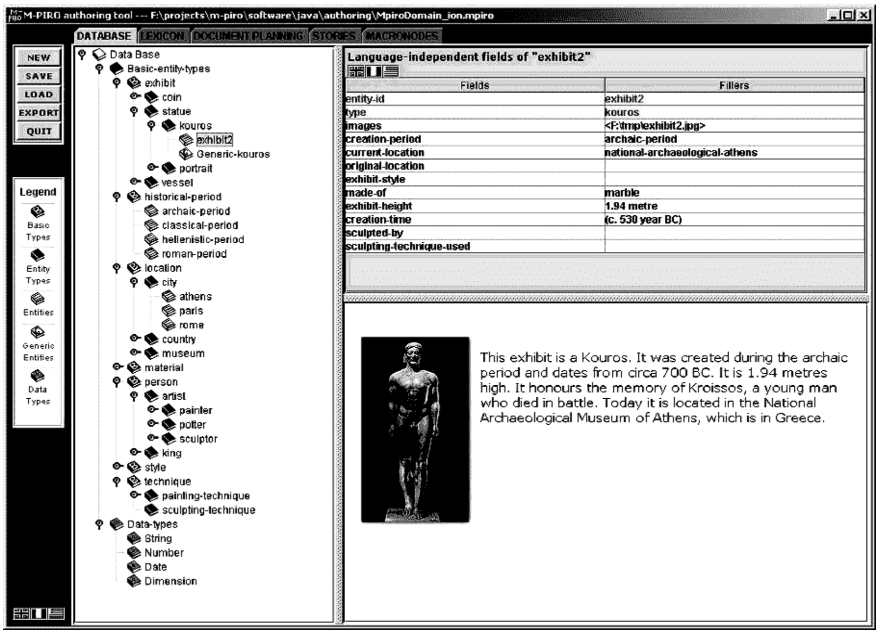

Fig. 6. Entering and previewing information about an object

To capture default information about all the entities of a type, *generic entities* can be introduced. For example, the *creation-period* field of the *Generic-kouros* entity in Figure 6 could be assigned the value *archaic-period*. This would indicate that, unless otherwise mentioned, a kouros belongs to the archaic period.

4 Clause Plans and Templates

During the authoring process, a micro-plan needs to be specified for each database field and language, to specify how the field can be expressed as a clause. Following Ilex, M-PIRO supports two forms of micro-plans: *clause plans* and *templates*.

In clause plans, the author specifies the verb to be used (from those available in the domain-specific lexicon, to be discussed in Section 5), the voice and tense of the

[2] Work on the integration of the authoring subsystem with the underlying generation engine is currently in progress. Figure 6 illustrates the previewing that will be available once the integration is complete.

resulting clause, the preposition, if any, to be included between the verb and the object, any desired adverb, and strings to be concatenated as adjuncts at the beginning or end of the clause. The micro-plan in Figure 5 leads to clauses like “This statue was sculpted by Polyklitus”. Appropriate referring expressions (e.g., “Polyklitus”, “a sculptor”, “him”) are generated automatically by the generation engine. Advanced clause-planning options allow the authors to select manually the case and type of a referring expression, the mood of the clause, and whether or not it can be aggregated. Clause-plans for the supported languages are often very similar, and verbs are kept aligned across the languages, as will be discussed in Section 5. The “get values from” buttons in Figure 5 speed up the authoring process by setting the fields of the clause plan to the same values as their counterparts in the other languages, where possible.

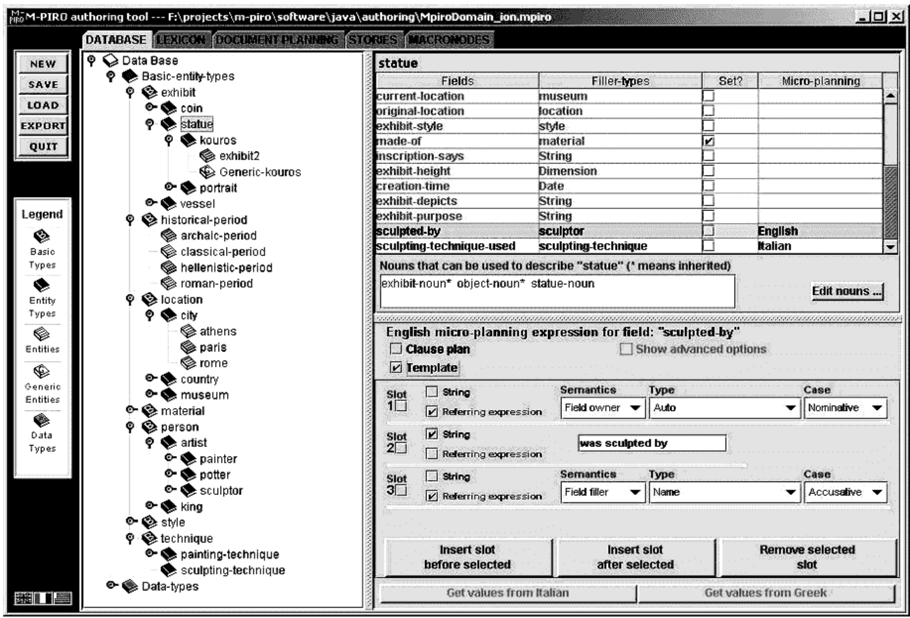

Fig. 7. A micro-plan in the form of a template

Templates provide stricter control over the surface form of the resulting clauses than do clause plans. A template is a sequence of slots, the values of which are simply concatenated to produce a clause. Figure 7 shows an alternative micro-plan for the *sculpted-by* field of Figure 5 in the form of a template. Each slot can be filled by a particular string, an expression referring to the owner of the field (the statue, in the case of *sculpted-by*), or a referring expression for the field’s filler (the sculptor). Templates carry less linguistic information than clause-plans, which does not allow the generation engine to exploit its full potential; for example, some forms of aggregation cannot be used with templates. However, templates are the only option when fields need to be rendered in forms other than clauses; e.g., copyright notes.

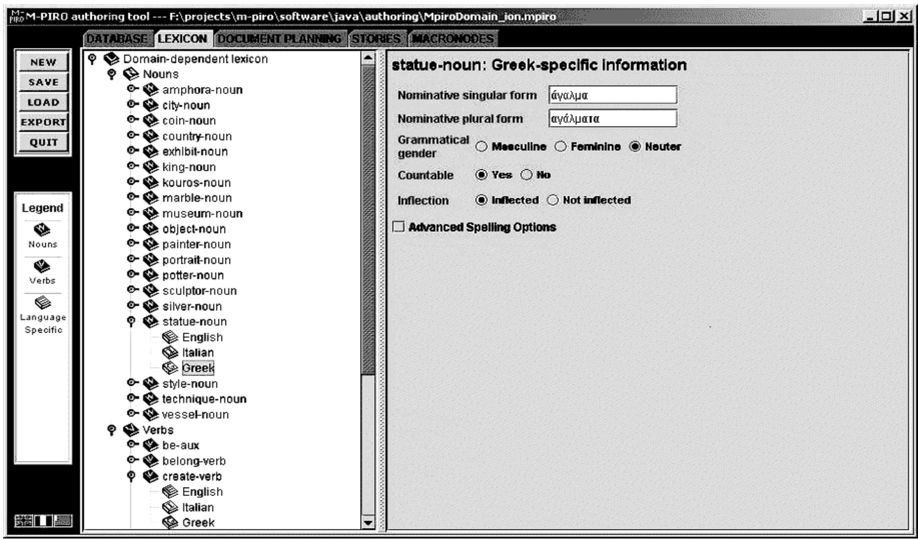

Fig. 8. Editing the domain-dependent lexicon

5 Domain-Dependent Lexicon

The domain-dependent lexicon contains entries for nouns and verbs, as shown in Figure 8. The entries for function words, such as articles and prepositions, are domain-independent and are kept separately. Nouns are associated with entity types; in Figure 7, the noun whose lexicon identifier is *statue-noun* is associated with the entity type *statue*, as can be seen in the area next to the "edit nouns" button. This licenses the generation engine to use *statue-noun* when referring to entities of this type (e.g., "this statue"). Additionally, each entity type inherits the nouns that have been associated with its super-types. In Figure 7, the entity type *statue* inherits the nouns *exhibit-noun* and *object-noun*, which have been associated with the type *exhibit*; hence, when referring to a statue, those nouns can also be used ("this exhibit" or "this object"). In practice, after defining the hierarchy of entity types, the author associates at least one noun with each entity type by selecting nouns from the domain-dependent lexicon.[3] If the domain-dependent lexicon does not contain the desired nouns, they first have to be inserted into the lexicon as shown in Figure 8. The system encourages the authors to keep the lexicons of the supported languages aligned by treating each entry as a triplet that contains nouns or verbs with equivalent senses in the three languages. For example, the entry of *statue-noun* contains "statue", "statua", and "άγαλμα", for English, Italian, and Greek. This helps maintain the same linguistic

[3] A few domain-independent noun entries also exist. They are linked to types of the Upper Model, and they are used when an entity type is not associated with any noun of the domain-dependent lexicon.

coverage across all languages. Entering verbs is similar, except that what leads the author to add a new verb is the need to use it in a clause-plan (Section 4).

Like most natural language generation systems, M-PIRO's domain-dependent lexicon is typically rather small; there are approximately 45 noun and 25 verb entries in the domain of the current web-based prototype, many of which (e.g., "amphora", "kouros") are unlikely to be found in general-purpose dictionaries. Hence, instead of attempting to reuse existing large-scale electronic dictionaries, we have opted for facilities that simplify entering new nouns and verbs. In the case of Greek nouns, for example, the dictionary of the underlying generation engine contains several features pertaining to the inflection pattern of the noun, the position of the stress in its various forms, etc. The authoring subsystem incorporates facilities that determine and add automatically these features by examining the nominative singular and plural forms of the noun. Morphology rules are also present, which generate automatically the remaining forms of the nouns, and similar facilities are available for verbs. The "advanced spelling options" button in Figure 8 allows those automatically generated forms to be inspected and corrected, if necessary.

6 Conclusions and Future Work

We have presented M-PIRO's authoring facilities, which help domain experts with no language technology expertise configure the system for new application domains. The authoring facilities currently allow the domain experts to manipulate the structure and content of the underlying database, as well as the domain-specific linguistic resources that are used during micro-planning and surface realization. Work is in progress to provide additional facilities for entering user types, stereotypes, and text planning schemata. Additional work is considering how support for *MacroNodes* can be provided; this is a technology deriving from the HIPS project [11] that allows canned texts to be customized according to the user model that has been activated, providing many of the benefits of full-scale generation.

Domain experts are currently using the authoring facilities to extend the domain of the web-based prototype, and the same facilities will be used to port M-PIRO's technology to another collection of exhibits in a virtual reality environment. Both activities will provide feedback on the usability of the authoring subsystem and the portability of the overall technology. A complementary strand of work is considering how existing museum databases can be interconnected with M-PIRO's components.

Finally, it would be interesting to examine how more active forms of previewing can be made available. Additional mark-up could be exploited to allow the authors to inspect database fields, micro-plans, or dictionary entries by selecting the corresponding clauses or words in the generated texts; this would help them repair anomalies in the content or realization of the texts. Mechanisms of this kind could be seen as an attempt to link symbolic authoring to the WYSIWYM approach [16], where authors interact with the system entirely via generated texts that reflect both the content of the database and the options that are available to update it.

References

1. I. Androutsopoulos, V. Kokkinaki, A. Dimitromanolaki, J. Calder, J. Oberlander and E. Not. "Generating Multilingual Personalized Descriptions of Museum Exhibits – The M-PIRO Project". Proc. *29th Conference on Computer Applications and Quantitative Methods in Archaeology*, Gotland, Sweden, 2001.
2. J. A. Bateman. "Enabling Technology for Multilingual Natural Language Generation". *Natural Language Engineering*, 3(1):15–55, 1997.
3. P. Cosi, F. Tesser, R. Gretter and C. Avesani. "Festival Speaks Italian!". Proc. *7th European Conference on Speech Communication and Technology*, Aalborg, Denmark, 2001.
4. R. Dale, S.J. Green, M. Milosavljevic, C. Paris, C. Verspoor and S. Williams. "Dynamic Document Delivery: Generating Natural Language Texts on Demand". Proc. *9th International Conference and Workshop on Expert Systems Applications*, Vienna, Austria, 1998.
5. A. Dimitromanolaki, I. Androutsopoulos and V. Karkaletsis. "A Large-Scale Systemic Functional Grammar of Greek". Presented at the *5th International Conference on Greek Linguistics*, Sorbonne, France, 2001.
6. M. Halliday. *An Introduction to Functional Grammar*. Edward Arnold, London, 1985.
7. E. Hovy. "Automated Discourse Generation Using Discourse Structure Relations". *Artificial Intelligence*, 63(1–2): 341–386, 1993.
8. W. Mann and S. Thompson. "Rhetorical Structure Theory: Towards a Functional Theory of Text Organization". *Text*, 3:243–281, 1988.
9. K.R. McKeown. *Text Generation*. Cambridge University Press, 1995.
10. M. Milosavljevic and J. Oberlander. "Dynamic Hypertext Catalogues: Helping Users to Help Themselves". Proc. *9th ACM Conference on Hypertext and Hypermedia*, Pittsburgh, PA, 1998.
11. E. Not and M. Zancanaro. "The MacroNode Approach: Mediating between Adaptive and Dynamic Hypermedia". Proc. *International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*, Trento, Italy, 2000.
12. J. Oberlander, M. O'Donnell, A. Knott and C. Mellish. "Conversation in the Museum: Experiments in Dynamic Hypermedia with the Intelligent Labelling Explorer". *New Review of Hypermedia and Multimedia*, 4:11–32, 1998.
13. M. O'Donnell, C. Mellish, J. Oberlander and A. Knott. "ILEX: An Architecture for a Dynamic Hypertext Generation System". *Natural Language Engineering*, 7(3):225–250, 2001.
14. C. Paris, K. Vander Linden, M. Fischer, A. Hartley, L. Pemberton, R. Power and D. Scott. "A Support Tool for Writing Multilingual Instructions". Proc. *14th International Joint Conference on Artificial Intelligence*, Montreal, Canada, pp. 1398–1404, 1995.
15. R. Power and N. Cavallotto. "Multilingual Generation of Administrative Forms". Proc. *8th International Workshop on Natural Language Generation*, pp. 17–19, Herstmonceux Castle, U.K., 1996.
16. R. Power and D. Scott. "Multilingual Authoring Using Feedback Texts". Proc. *36th Meeting of the ACL and 17th International Conference on Computational Linguistics*, pp. 1053–1059, Montreal, Canada, 1998.
17. E. Reiter and R. Dale. *Building Applied Natural Language Generation Systems*. Cambridge University Press, 2000.
18. M. Theune, E. Klabbers, J.R. De Pijper, E. Krahmer and J. Odijk. "From Data to Speech: A General Approach". *Natural Language Engineering*, 7(1):47–86, 2001.
19. G. Xydas and G. Kouroupetroglou. "Text-to-Speech Scripting Interface for Appropriate Vocalization of e-Texts". Proc. *7th European Conference on Speech Communication and Technology*, Aalborg, Denmark, 2001.