Revista de Lenguas para Fines Específicos 23.2 (2017), pp. 86-110
http://dx.doi.org/10.20420/rlfe.2017.174
ISSN: 2340-8561

Servicio de Publicaciones y Difusión Científica
Universidad de Las Palmas de Gran Canaria

The process of building the upper-level hierarchy for the aircraft structure ontology to be integrated in FunGramKB

Ángel Felices-Lago[1]
University of Granada
Ángela Alameda-Hernández
University of Granada

ABSTRACT

In this article we collect a corpus of texts which operate with a controlled language (ASD Simplified Technical English) in order to facilitate the development of a new domain-specific ontology (the aircraft structure) based on a technical discipline (aeronautical engineering) included in the so called "hard" sciences. This new repository should be compatible with the Core Ontology and the corresponding English Lexicon in FunGramKB (a multipurpose lexico-conceptual knowledge base for natural language processing (NLP)), and, in the same vein, should eventually give support to aircraft maintenance management systems. By contrast, in previous approaches we applied a stepwise methodology for the construction of a domain-specific subontology compatible with FunGramKB systems in criminal law, but the high occurrence of terminological banalisation and the scarce number of specific terms, due to the social nature of the discipline, were added problems to the most common NLP difficulties (polysemy and ambiguity). Taking into consideration previous results and the complexity of this task, here we only intend to take the first step towards the modelling of the aircraft ontology: the development of its taxonomic hierarchy. Consequently, the hierarchy starts with the whole system (i.e., an aircraft) and follows the traditional decomposition of the system down to the elementary components (top-down approach). At the same time,

1 **Corresponding author** – Departamento de Filología inglesa y alemana, Facultad de Ciencias Económicas y Empresariales, Campus Universitario de Cartuja, s/n, 18071, Granada, Spain.

Email: afelices@ugr.es

we have collected a corpus of 2,480 files of aircraft maintenance instructions, courtesy of *Airbus* in Seville. For the bottom-up approach (under construction), we consult specialised references end explore the corpus through the identification and extraction of term candidates with DEXTER, an online multilingual workbench especially designed for the discovery and extraction of terms.

Keywords: FunGramKB, terminology, ontology building, taxonomic hierarchy, aircraft ontology.

1. Introduction

Ontology comes from Greek philosophy and deals with the study of being and existence and how existing things relate to each other. This concept has been incorporated to use in computer science. In the context of the Semantic Web, the most commonly used definition comes from Studer et al. (1998): "An ontology is a formal explicit specification of a shared conceptualization of a domain of interest". He emphasizes the automated processing, the consensus on the contents and the orientation towards a specific domain.

An advantage of ontologies is the robust reasoning service that they incorporate from their formal foundation. Reasoning services mainly support the activities of creating a knowledge representation for a domain and the retrieval of knowledge from it. In this study we think that a stepwise methodology for the construction of domain-specific ontologies can be compatible with the reasoning services offered by FunGramKB (Periñán-Pascual & Arcas-Túnez, 2010a, 2014), which is a multipurpose lexico-conceptual knowledge base for natural language processing (NLP) systems. Consequently, we developed a three-phase method for this type of ontology building: corpus collection, terminological work (automatic extraction and manual filtering) and conceptual modelling tasks (conceptualisation, hierarchisation and subsumption). This methodology was initially applied to the *Globalcrimeterm* subontology (Felices-Lago & Ureña-Gómez-Moreno, 2012, 2014; Carrión-Delgado & Felices-Lago, 2014; Felices-Lago 2015; Alameda-Hernández & Felices-Lago, 2016). However, the high occurrence of terminological banalisation as well as the social nature of this criminal law subdomain highlighted the limitations for building a consistent domain-specific ontology (Felices-Lago, 2016).

To overcome this problem and other common NLP difficulties (polysemy and ambiguity) which have hindered the construction of the aforementioned subontology, we intend to explore two different solutions here: (i) to develop a new domain-specific ontology based on a more technical discipline which may eventually give support to aircraft maintenance management systems; (ii) to operate with a well-known English-based controlled language, ASD Simplified Technical English (ASD-STE) (Wojcik, Holmback & Hoard 1998, Møller & Christoffersen, 2006), and make it compatible with the Core Ontology and the corresponding English lexicon in FunGramKB. Taking into consideration the complexity of both tasks, in this article we intend to take the first step towards the modelling of an aircraft ontology which is based on the development of its taxonomic hierarchy.

For reaching this purpose, in sections 2 and 3 we explore the *state-of-the-art* of aeronautical ontology-building (domain, task or application ontologies) and also refer to the advantages of working with documents written in a controlled language such as the ASD Simplified Technical English (ASD-STE). In section 4 we summarise the main characteristics of the knowledge base named FunGramKB, with special attention to the FunGramKB Core Ontology and its capability to include specialised knowledge by establishing links to satellite or domain-specific ontologies. In section 5, we explain the guidelines and the steps taken for the building of the aircraft taxonomic hierarchy. In section 6 we offer the preliminary and schematic results of the top-down analysis and present the upper-level hierarchy of this ontology. Section 7 offers some conclusions.

2. Aeronautical ontology-building: state of the art

According to Studer et al. (2007), ontologies can be categorised into the following types: (i) *top level ontologies* that cover general and abstract concepts which can be reused in other ontologies (i.e. notions of time or space); (ii) *domain or task ontologies* which cover knowledge about a specific domain (i.e. the aircraft) or a general task (i.e. cooking). As a further specialisation, there are (iii) *application ontologies*, which can cover and refine specific aspects of domain ontologies for use in that specific application and with certain usage scenarios in mind. A commonly used language is Ontology Web Language (OWL), which has been developed to be compatible to the World Wide Web (WWW) architecture.

The ontology we intend to develop in the context of this article (the aircraft structure ontology) can be categorised as a *domain ontology* and, as far as we know, has not been created so far with the unique purpose of formally representing the taxonomic hierarchy of the prototypical aircraft structure. However, in aeronautics, formal and explicit specifications of shared conceptualisations have existed for a long time. An example is geographical navigation, based on the cardinal points together with meridian and latitude coordinates. Conceptualisation enables sharing references among the aeronautic community, but also between human actors and artifacts. An example is the electronic maps coupled with Global Positioning Systems (Reiss et al., 2006). Another example of conceptualisation shared by humans and machines is the explicit conception and application of domain procedures. Some are described in regulations: navigation procedures (Keller, 2016), and others cannot be modified:

this is the case for safety relevant procedures (Garst, 2009). The task that the flight crews must accomplish is also strongly related to procedures. Reiss et al. (2006) refer to the phases-of-flight described by Travers (2000) to enable a user-oriented description of the fragmented sequential phases a flight crew must go through for the accomplishment of a mission. However, the amount of existing taxonomies and ontologies in aeronautics is extremely diverse, as it is the case of route planning (Liu et al., 2009, Niaraki & Kim, 2009), traffic flow management (van Putten et al., 2008), flight simulator fidelity evaluation (Durak et al., 2014), aircraft assembly (Hongjung et al., 2007), aircraft fault knowledge (Yi et al., 2009, Zhou & Li, 2011) or maintenance task support (Brusilovsky & Cooper, 2002, Wu et al., 2014), among others.

None of the aforementioned ontologies conceptualises the aircraft structure to include the most specific details, except for the *ontology for aircraft design* proposed by Ast, Glas & Roehm (2013), and only to a limited extent[2]. They intend to describe the development of the aircraft ontology following the NeOn process model, in particular the experiences from applying the NeOn methodology to the resulting aircraft ontology. This ontology is also an OWL ontology that covers system decomposition and component parameters of a single aisle civil transport aircraft. However, the project is based on the Diploma thesis of Markus Ast (2012), which models the physical structure of a typical passenger aircraft like an *Airbus A320* on a high abstraction level. Ast (2012, p. 67) claims that "this abstraction level can, at any time, be extended with more details as the need arises". But it is also a fact that, according to him, aircraft structural concerns are but one of several views that are relevant for an aircraft and, consequently, his contribution only includes the most basic and prominent parts or sections of the aircraft structure. Most authors in the field refer to the *Air Transportation Association* (ATA) to categorise the parts of an aircraft, and the ATA numbers have become an accepted categorisation of aircraft parts with reference to aircraft systems. However, we think that it is possible to take a step forward and ontologically conceptualise in depth the parts and components of a plane like the *Airbus A320* and reach a detailed hierarchisation of the relevant concepts. The results of this process will be shown in section 6.

3. The ASD Simplified Technical English (ASD-STE)

As stated in section 1, the development of a new domain-specific ontology, compatible with the Core Ontology in FunGramKB and, at the same time, able to

2 This *aircraft ontology* contains 96 classes and 224 object properties. For demonstration purpose the ontology also contains 22 individuals. See Ast (2012) for an in-depth description of the ontology.

overcome NLP difficulties such as polysemy and ambiguity, inspired our proposal to build the aircraft structure ontology. This artefact repository is based on aeronautical engineering, a technological discipline which gave birth to the controlled language known as ASD-STE (Aero-Space and Defence Industries Association of Europe - Simplified Technical English) (ASD 2013). This is a controlled language developed in the early Eighties to help the users of English language maintenance documentation understand what they read and avoid misunderstandings. Originally inspired by a language called ILSAM[3] (Adriaens & Schreors, 1992), the language had its origins in 1979, but it was only in 1986 when it was officially presented for the first time, then under the name AECMA[4] Simplified English. It received its current name in 2004 when AECMA merged with two other associations to form ASD.

According to Kuhn (2014, p. 136), "the main purpose of this language is to make texts easier to understand, especially for non-native speakers". Today, this language is maintained by the Simplified Technical English Maintenance Group. ASD-STE is based on English with restrictions expressed in about 60 writing rules (Part 1) and a dictionary of controlled vocabulary (Part 2). The writing rules cover aspects of grammar and style and the dictionary specifies the general words that can be used. These words were chosen for their simplicity: "One word - one meaning". Usually, each word is permitted for only one part of speech. For example, the word 'oil' is specified as a noun. Therefore, the word 'oil' must not be used as a verb. There is a fixed vocabulary consisting of terms common to the aerospace domain. Additionally, user-defined "Technical Names" and "Technical Verbs" can be introduced. In consequence, Kuhn maintains (2014, p. 136) that, "even though its restrictions make ASD-STE considerably more precise than full English, it does not allow for reliable automatic interpretation. Full expressiveness and full naturalness of unconstrained English are retained, but also its complexity".

Taking into consideration the objections raised by Kuhn with reference to the feasibility of ASD-STE to carry out a reliable automatic language processing, we think instead that there is a considerable compatibility between the "general" words from the ASD-STE 100 Dictionary of controlled vocabulary and the FunGramKB Core Ontology conceptual units, but this is something that can only be tested in the final phase of the aircraft structure ontology building, once FunGramKB`s reasoning engine is applied.

---

[3] International Language of Service and Maintenance.

[4] Association Européenne des Constructeurs de Matériel Aérospatial.

4. The architecture of FunGramKB

FunGramKB is an online environment for the semiautomatic construction of a multipurpose lexico-conceptual knowledge base for NLP systems, and more particularly for natural language understanding (Periñán-Pascual & Arcas-Túnez, 2004, 2007, 2010a, 2010b; Periñán-Pascual & Mairal-Usón, 2009, 2010). As can be observed in Figure 1 below, FunGramKB comprises three major knowledge levels, consisting of several independent but interrelated modules:

(1) Lexical level: The *Lexicon* stores morphosyntactic, pragmatic and collocational information about lexical units and the *Morphicon* handles cases of inflectional morphology.

(2) Grammatical level: The *Grammaticon* stores the constructional schemata which help to construct the semantics-to-syntax linking algorithm (Van Valin & LaPolla, 1997, Van Valin, 2005).

(3) Conceptual level: The *Ontology* is presented as a hierarchical catalogue of the concepts that a person has in mind, so here is where semantic knowledge [5] is stored in the form of Meaning Postulates (MPs), a group of logically connected predications which articulate the generic features of a concept and are written in a conceptual representation language known as COREL (Periñán-Pascual & Mairal-Usón, 2010). The ontology also consists of a general-purpose module (i.e. Core Ontology) and several domain-specific terminological modules (i.e. satellite ontologies or subontologies)[6]. The *Cognicon* stores procedural knowledge by means of scripts, that is, conceptual schemata in which a sequence of stereotypical actions is organised on the basis of temporal continuity, and more particularly on Allen's temporal model (Allen 1983; Allen & Ferguson, 1994); e.g. 'dine in a restaurant', 'pilot a plane', 'pay with a credit card', etc. The *Onomasticon* stores information about instances of entities and events such as 'Taj Mahal', 'September 11', 'Donald Trump', etc.: episodic knowledge. This module stores two different types of schemata (i.e. snapshots and stories), since instances can be portrayed synchronically or diachronically.

---

[5] The underlined types of knowledge follow the distinctions established within the framework of cognitive psychology.

[6] The aircraft structure ontology could be one of the satellite ontologies linked to the Core Ontology in FunGramKB.

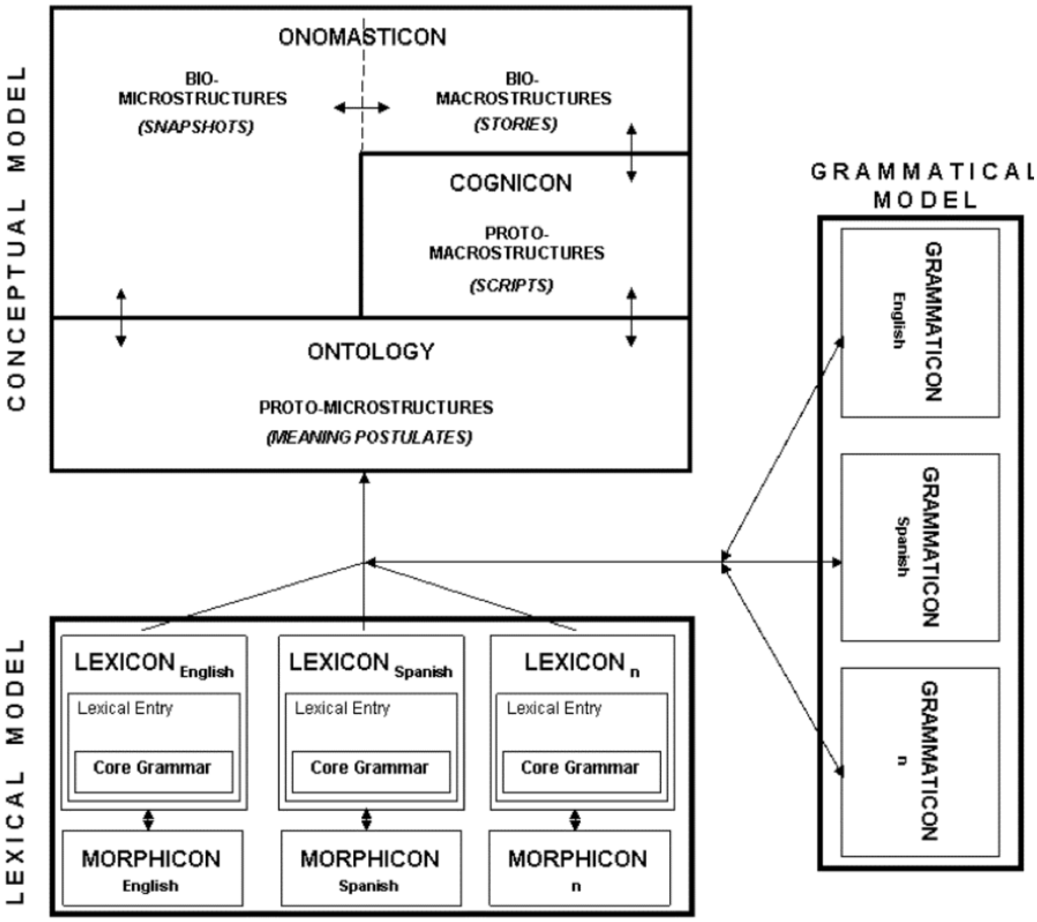


Figure 1. FunGramKB modules.

Besides that, FunGramKB Ontology is a conceptual taxonomy, derived from linguistic concepts, in which interlinguistic differences in syntactic constructions do not involve conceptual differences. It is general-purpose, and not domain-specific. However, since expert knowledge stems from general knowledge, it can be extended to include specialized knowledge by establishing links to satellite or domain-specific ontologies, be it medicine, law, chemistry or aeronautical engineering. The concepts of FunGramKB belong to three levels. The upper level is composed of 42 metaconcepts, marked with the symbol #. They constitute the upper level in the taxonomy as a result of the analysis of the most relevant linguistic ontologies, such as DOLCE (Gangemi et al., 2005), SIMPLE (Lenci, 2008), SUMO (Niles & Pease, 2001), etc. These metaconcepts are distributed in three subontologies: #ENTITY, #EVENT, and #QUALITY. Second level concepts (immediately under metaconcepts) are marked by the sign + (e.g. +AIRCRAFT_00). These concepts are used in the meaning postulates that define basic and terminal concepts, and also encode the selection restrictions in thematic frames. The third level is composed of terminal concepts, marked by $ (e.g. $RUDDER_00). The difference between basic and terminal concepts is that basic concepts are used to define other concepts in Meaning Postulates, whereas terminal concepts are not. Obviously, in the satellite ontologies for specialized knowledge, terminal concepts in FunGramKB will have to be promoted to basic concepts under the adequate circumstances, or some basic concepts in the Core Ontology should become mirror concepts to cover the expert knowledge (Felices-Lago, 2016).

Consequently, we follow the principle that ontologies, like the ones included in FunGramKB, consist of metaconcepts and concepts and the concepts are organized in a hierarchical structure formed by IS-A relations between these concepts. In this context, the hierarchy we intend to conceptualize starts with the whole system (i.e., an aircraft) and follows traditional decomposition of the system down to the elementary components (top-down approach). It also uses COREL interface language, distinguishes three different conceptual levels (each one of them with concepts of a different type: Metaconcepts (#), basic concepts (+) or terminals ($)), and allows multiple non-monotonic inheritance.

5. Methodological steps to build the upper-level hierarchy of the ontology

As explained above, the aircraft structure ontology can be defined as a hierarchical taxonomy of specialised concepts belonging to an expert area of knowledge: basically, aeronautical engineering. It thus serves the purpose of enhancing FunGramKB with specialised knowledge, considering that this knowledge base was originally implemented to work with elementary common-sense concepts of human cognition. The Core Ontology and the domain-specific (or satellite) ontologies might eventually become connected as shown in Figure 2:

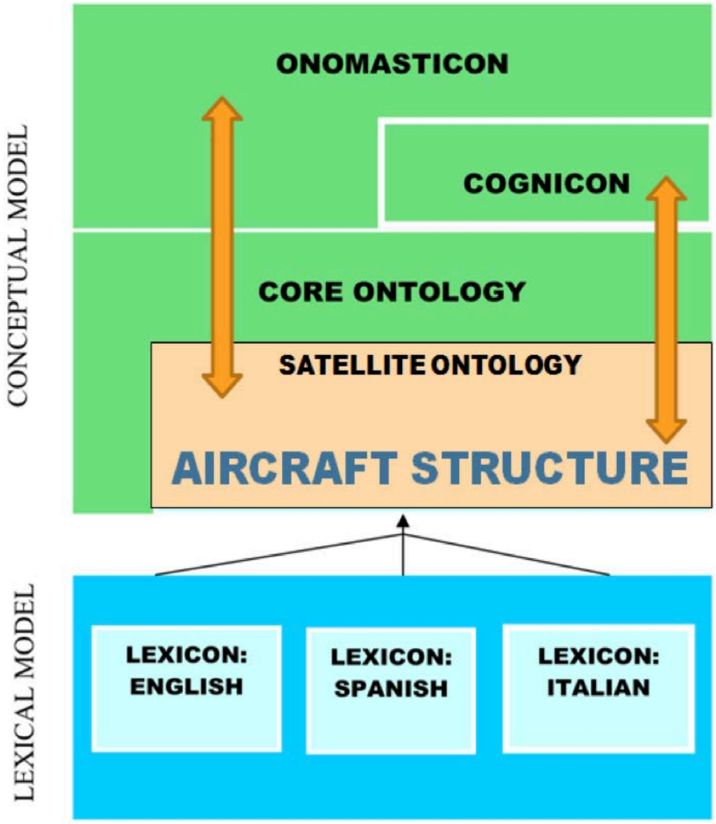


Figure 2. Extension of FunGramKB architecture including Satellite Ontologies.

In order to feed the domain-specific ontology and the corresponding lexica, we intend to take the following steps:

(1) The organisation of the core conceptual structure of the aircraft structure is carried out by analysing the evidence obtained in the epistemological frameworks of previous aeronautical ontologies (see section 2 and appendix 1) and also by seeking expert advice through professionals as well as by consulting reliable reference sources such as technical handbooks, specialised dictionaries and glossaries (see appendix 2). These terminological units, based on the deductive approach to the thematic domain, will be verified or substituted through the inductive methodological phase of the ontology building (bottom-up approach, under construction). Thus, the field work will either ratify or not ratify the coherence and reliability of the structure initially proposed. This process is known in the practice of ontology design as 'common sense'.

(2) The defining words of the thematic domain in the field of aircraft structure must be also identified through intensive searches in the controlled ASD-STE corpus of

2480 files[7] written in the ASD-STE controlled language at our disposal, courtesy of *Airbus* in Seville (see appendix 3). In this way defining texts can be processed more easily and automatically. For the exploration of this corpus and the identification and extraction of term candidates (i.e. unigrams, bigrams and trigrams) we use the new complementary FunGramKB tools for NLP: DEXTER, an online multilingual workbench especially designed for the discovery and extraction of terms,8 and DAMIEN, a workbench that allows researchers to do text analytics by integrating corpus-based processing with statistical analysis and machine-learning models for data mining tasks (Periñán-Pascual, 2015; Periñán-Pascual & Mestre-Mestre, 2015; Periñán-Pascual (2017, this volume).

(3) The information of each concept must be included in FunGramKB through its online editor, which will connect, on the conceptual level, the satellite ontology which we will have created with the Core Ontology, the cognicon and the onomasticon; and, on the lexical level, with the lexicon corresponding to the language selected for this study, as can be seen in Figure 3:

7 This repository has been entitled *The Airbus corpus* for our purposes, and its main characteristics are described in appendix 3.

8 The suitability of the metric used in DEXTER is based on three fundamental notions: (i) *Salience*, i.e. which indicates the uniqueness or prevalence of a term in the data collection; (ii) *Relevance*, i.e. which measures the tendency of term usage between a domain-specific corpus and a general-purpose one, and (iii) *Cohesion*, i.e. which quantifies the degree of stability of multi-word terms. SRC can be described as a hybrid method, not only because it combines the linguistic approach with the statistical one, but also because it combines an *AKE (K: KEY) measure (i.e. salience) with **ATE measures (i.e. relevance and cohesion): *AKE stands for automatic keyword extraction and **ATE refers to automatic term extraction.

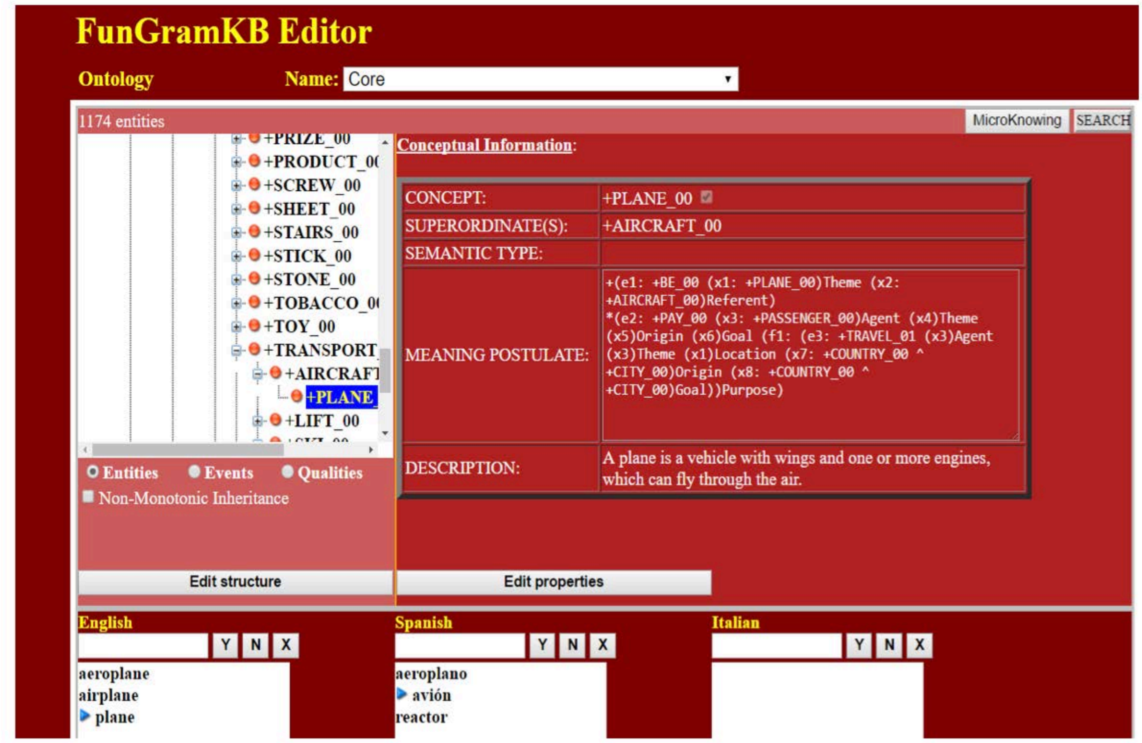


Figure 3. The concept +PLANE_00 and its integration in the Ontology of FunGramKB editor.

(4) A specific dictionary-like interface for this ontology must be designed and its hierarchical information must be able to be accessed both by humans and by the machine through the language of conceptual representation COREL, as can be seen in Figure 4:

Figure 4. The conceptual path of +AIRCRAFT_00 in the Core Ontology.

(5) The ontology must be semi-automatically[9] populated in its hierarchical structuring: appropriate meaning postulates must be constructed and domain-related basic concepts, terminal concepts must be created (currently, under construction).

A further step to be taken in the future includes:

(6) Specialised lexica corresponding to the English and other languages must be semi-automatically populated with their pertinent lexical information.

6. Results

Taking into account the strict application of the first two steps described above and combining term automatic extraction (DEXTER) and manual filtering, the proposed schematic results of the upper-level hierarchy for the aircraft structure ontology[10] are shown as follows:

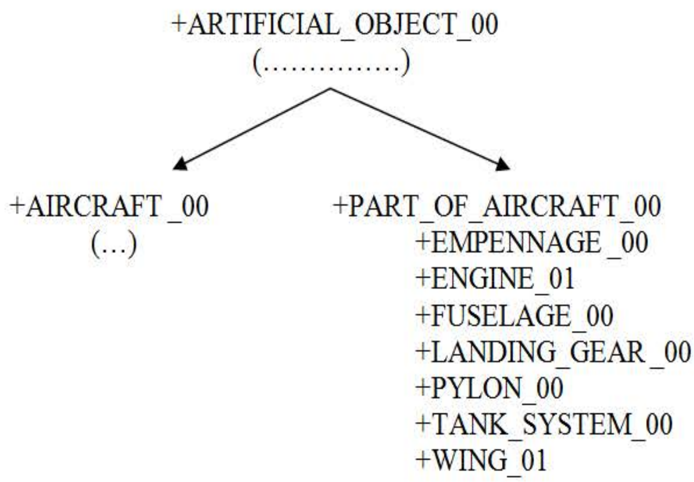


Figure 5. The upper-level hierarchy of the aircraft structure.

---

9 Semi-automatic population means that the candidate terms referred to in steps 5 and 6 have been automatically selected through DEXTER Term Extractor. After that, the terminologists and/or linguists have done the final selection of the winning candidates by checking lexicographical or other specialised resources and, at the same time, producing the MPs of the selected concepts.

10 The selected aircraft prototype for the building of this ontology is the *Airbus 320*, one of the most common commercial passenger twin-engine jet airliners manufactured by *Airbus.*

All the top level hierarchical concepts selected in this paper belong to the #ENTITY subontology and stand for nouns. The "part of aircraft" subordinate concepts (empennage, engine, fuselage, landing gear, pylon, tank system and wing) include a total of 36 basic concepts and 62 terminal concepts. The hierarchy in the specific-domain ontology should start with the whole system (i.e., an aircraft), but meronymy is not a permitted ontological relation in FunGramKB and, consequently, the aircraft parts or components are concepts branching off from a different conceptual path in the Ontology, but sharing with "aircraft" a common superordinate concept: +ARTIFICIAL-OBJECT_00. Following the traditional decomposition of the system down to the elementary components (top-down approach), the remaining concepts in the Ontology are eventually subordinate to at least one of the diverse parts of the aircraft, as can be observed in the schematic examples below.

(1) Part of the aircraft: engine

The engine is the most complex and detailed subcomponent of the aircraft. As can be observed below, we collect the conceptual terms representing the diverse parts of one of the most prototypical engines: the turbofan, which is generally wing-mounted. The turbofan[11] or fanjet is a type of airbreathing jet engine that is widely used in aircraft propulsion.

+PART_OF_AIRCRAFT_00
    +ENGINE_01
        +WING_MOUNTED_ENGINE_00
            +TURBOFAN_00 +PART_OF_TURBOFAN_00
                (...)
        (...)
+PART_OF_TURBOFAN_00
    $COMBUSTION_CHAMBER_00
    +FAN_00
+PART_OF_FAN_00
    $BOOSTER_SPOOL_00
    $FAN_BLADE_00
    $FAN_CASE_00
    $FAN_DISK_00
    $FAN_HUB_FRAME_00
    $LEVER_ARM_00

11 This definition can be found in *Wikipedia* (https://en.wikipedia.org/wiki/Turbofan [20/01/2017]).

$NOSE_LIP_00
+HIGH_PRESSURE_COMPRESSOR_00
+PART_OF_HIGH_PRESSURE_COMPRESSOR_00
$COMPRESSOR_CASE_00
$COMPRESSOR_SHAFT_00
$HP_COMPRESSOR_STATOR_00
+HIGH_PRESSURE_TURBINE_00
+PART_OF_HIGH_PRESSURE_TURBINE_00
$HP_TURBINE_CASE_00
$HP_TURBINE_SHAFT_00
$TURBINE_CENTER_FRAME_00
+LOW_PRESSURE_COMPRESSOR_00
+PART_OF_LOW_PRESSURE_COMPRESSOR_00
$INTEGRATED_BLADED_ROTOR_00
+LOW_PRESSURE_TURBINE_00
+PART_OF_LOW_PRESSURE_TURBINE_00
$LP_TURBINE_BLADE_00
$LP_TURBINE_CASE_00
$TURBINE_DISK_00
+NACELLE_00
+PART_OF_NACELLE_00
$EXHAUST_CONE_00
$EXHAUST_NOZZLE_00
$FAN_COWL_00
$INLET_00
$THRUST_REVERSER_00

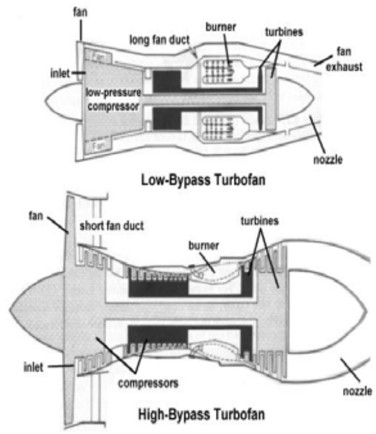


Figure 6. Models of turbofan engines[12].

(2) Part of the aircraft: empennage

The empennage includes the entire tail group and consists of fixed surfaces such as the vertical stabilizer and the horizontal stabilizer.[13] The movable surfaces include the rudder, the elevator, and one or more trim tabs.

+PART_OF_AIRCRAFT_00

---

12 All the pictures in figures 6 to 12 have been taken from *Google* and *Wikipedia* and we have also checked that their use and publication is not restricted, particularly for teaching o academic purposes.

13 This definition can be found in *Wikipedia* (https://en.wikipedia.org/wiki/Empennage [20/1/2017]).

+EMPENNAGE _00
+PART_OF_EMPENNAGE_00
$ELEVATOR_00
$HORIZONTAL_STABILIZER_00
$RUDDER_00
$STABILATOR_00
$TAIL_ CONE_00
$TRIM_TAB_00
$VERTICAL_STABILIZER_00

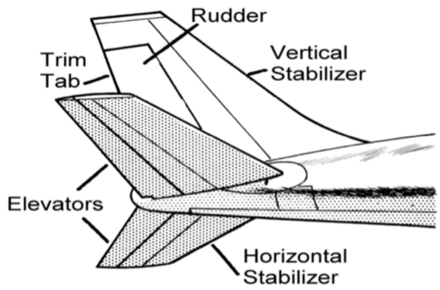


Figure 7. Model of empennage.

(3) Part of the aircraft: wing

The wing is the part of a heavier-than-air aircraft that produces aerodynamic lift to support the aircraft in flight and counters the forces of gravity (Kumar & Marshall 2005, p. 706). Usually, aircraft wings have various devices, such as flaps or slats that the pilot uses to modify the shape and surface area of the wing to change its operating characteristics in flight.

+PART_OF_AIRCRAFT_00
+WING_01
+PART_OF_WING_01
$AILERON_00
$FLAP_00
$SLAT_00
$SPOILER_00
$WINGLET_00

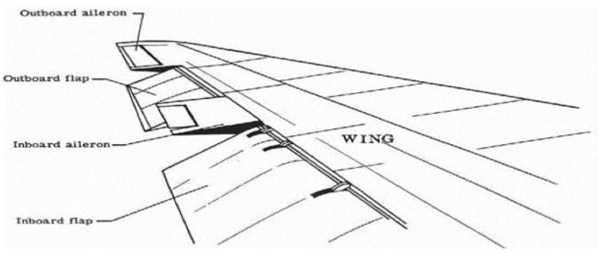


Figure 8. Model of aircraft wing.

(4) Part of the aircraft: fuselage

It is well known that the fuselage is the central body of a plane, to which the wings and tail assembly are attached and which accommodates the crew, passengers, and cargo.

+PART_OF_AIRCRAFT_00
+FUSELAGE_00
+PART_OF_FUSELAGE_00
+AFT_NON_CONSTANT_SEC TION
+CONSTANT_SECTION_00
+FORWARD_NON_CONSTANT_SECTION_00
+CABIN_00

Figure 9. Model of fuselage.

+PART_OF_CABIN_00
    $CARGO_COMPARTMENT_00
    $COCKPIT_00
    +PASSENGER _COMPARTMENT_00
+PART_OF_PASSENGER _COMPARTMENT_00
    $ACCESSORY_COMPARTMENT_00

$BUSINESS_CLASS_COMPARTMENT_00

    $ECONOMY_CLASS_COMPARTMENT_00
    $FIRST_CLASS_COMPARTMENT_00
    $GALLEY_00
    $LAVATORY_00

(5) Part of the aircraft: landing gear

This is the part of an airplane structure that supports the aircraft when it is not flying and permits takeoffs and landings without damage (Kumar & Marshall, 2005, p. 391) It generally uses wheels.

+PART_OF_AIRCRAFT_00
    +LANDING_GEAR _00
+PART_OF_LANDING_GEAR _00
    $ACTUATING_CYLINDER_00
    $BRAKE_BAR _00
    $DOWNLOCK_ACTUATOR_00
    $LOCK_STAY_00
    $PITCH_TRIMMER_00
    $SHOCK-ABSORBER_00
    $STRUT_ASSEMBLY_00
    $TORQUE_LINK_00

Figure 10. Model of landing gear.

(6) Part of the aircraft: pylon

The pylon is the structure that holds a pod or an engine nacelle to the wing or fuselage (Kumar & Marshall, 2005, p. 514).

+PART_OF_AIRCRAFT_00
    +PYLON_00
+PART_OF_PYLON_00
    $AFT_ENGINE_MOUNT_00

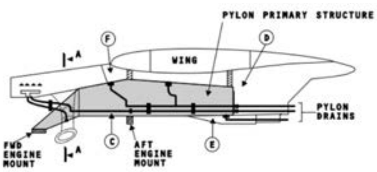


Figure 11. Model of pylon.

$FWD_ENGINE_MOUNT_00
$PYLON_DRAIN_00

(7) Part of the aircraft: tank system

Aircrafts include diverse tank systems for waste and liquids such as fuel, oil or potable water. However, fuel is one of the most crucial aspects in a plane due to its volume and consumption. Fuel tanks have to store a lot of weight and are generally distributed in diverse areas of the plane to facilitate flight stability and safety.

+PART_OF_AIRCRAFT_00
+TANK_SYSTEM_00
$OIL_TANK_00
$POTABLE_WATER_TANK_00
$WASTE_TANK_00
+FUEL_TANK_00
+PART_OF_FUEL_TANK_00
$CENTRE_TANK_00
$INNER_TANK_00
$OUTER_TANK_00
$TRIM_TANK_00
$VENT_TANK_00

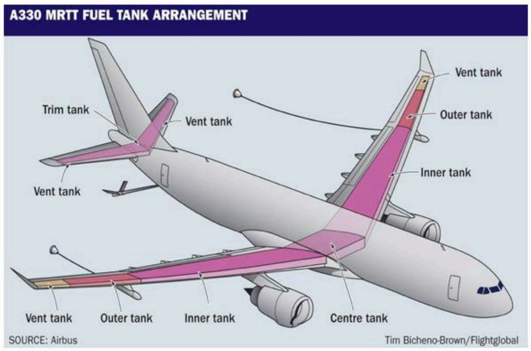


Figure 12. Model of distribution of fuel tanks.

7. Concluding remarks

Previous attempts to develop a domain-specific ontology based on deep semantics and linked to the Core Ontology of FunGramKB have experienced countless difficulties, particularly when dealing with an efficient reasoning for NLP tasks. The possibility to improve these results in the future requires a new strategy founded in two methodological choices: (1) to work with a controlled natural language which avoids the most common problems of NLP such as polysemy and ambiguity and, at the same time, to operate with the contents of a scientific discipline which may offer consistent methodological rigor, exactitude, and objectivity. Consequently, in this work we have taken the first step toward that direction.

We have aimed at the initial stages for the building of a domain-specific ontology (the aircraft structure), which is grounded on a technology-based scientific field, as is the case of aeronautical engineering. We also have chosen a set of documents for corpus compilation, *the Airbus Corpus*, which has been produced under the lexical and syntactic restrictions established by the ASD Simplified Technical English. This aircraft maintenance repository has been enriched with the exhaustive look up of specialised lexicographical and technical sources dealing with aircraft design and maintenance. The application of DEXTER (the FunGramKB Suite term extractor) on the sampled texts has rendered 6,966 n-grams (unigrams, bigrams and trigrams). The manual filtering of the candidates has helped us to make a previous selection of the concept-related terms which make up the the upper-level hierarchy for the aircraft structure ontology to be integrated in FunGramKB and shown in section 6. However, in order to conclude the construction of this ontology there are tasks which have to be accomplished in the near future: (i) to include the information of each concept in FunGramKB through its online editor; (ii) to construct appropriate meaning postulates for the domain-specific basic and terminal concepts according to COREL interface language; (iii) to populate the specialised lexica with their pertinent lexical information, corresponding to the English language and other languages which might be incorporated.

It is also relevant to note that the *ontology for aircraft design* proposed by Ast (2012) only contains the most basic and prominent parts or sections of the aircraft structure and leaves open the possibility for further specification and hierarchisation. Our proposal extends considerably the ontological account and the quantity of the aircraft parts collected by Ast and, in the same vein, intends to offer a well-grounded method to categorise some sections of the *Air Transportation Association* (ATA) specifications. This method includes FunGramKB deep semantics principles, particularly conceptual subsumption and multiple non-monotonic inheritance and, as a consequence, consistent results can only be attained once the phases involving the specialised conceptual modelling and the population of the lexica are completed. Moreover, the final product might offer a repository on the aircraft structure for its potential exploitation in tasks of automatic translation, retrieval of information, aircraft maintenance management systems and aircraft design.

About the authors

Ángel Felices-Lago is Full Professor and works as a member of the Department of English and German Philology (University of Granada, Spain) since 1984. He teaches English and Spanish for business and tourism. His main areas of research interest go from lexicology, discourse analysis and axiological linguistics to NLP applied to LSP. He has co-authored or co-edited 10 scholar and pedagogical books or textbooks and has also published over 80 scholarly articles and reviews in specialized national and international journals and volumes. He has served as invited reviewer or member on the editorial and scientific boards of a dozen

journals and has also led or taken part in various international academic projects funded the Spanish Ministry of Education or the European Union (Tempus, Leonardo, Erasmus, Erasmus-Mundus, etc.).

Ángela Alameda-Hernández is Assistant Lecturer at the Department of English and German Philology, University of Granada (Spain), where she teaches English Grammar and English culture. Her main research interests are discourse analysis, terminology and NLP. She has extensively published academic articles, book chapters and a textbook on these topics, including media discourse, specialized terminology and on foreign language didactics. She has taken part in a number of pedagogical innovation projects and research projects funded both by the University of Granada and by the Spanish Ministry of Education.

Acknowledgements

This article is based on research carried out within the framework of the Project FFI2014-53788-C3-1-P, which is funded by the Spanish Ministry of Economy and Competitiveness.

Article history

Paper received: 19 February 2017
Paper received in revised form and accepted for publication: 27 April 2017

References

Adriaens, G., & Schreors, D. (1992). From COGRAM to ALCOGRAM: Toward a controlled English grammar checker. In *Proceedings of the 14th conference on Computational Linguistics-Volume 2* (pp. 595-601). Association for Computational Linguistics.

Alameda-Hernández, Á. & Felices-Lago, Á. (2016). The integration of the concept +CRIME_00 in FunGramKB and the conceptualization or hierarchization problems involved. In C. Periñán-Pascual & E. Mestre-Mestre (Eds.) *Understanding meaning and knowledge representation: From theoretical and cognitive linguistics to natural language processing* (pp. 319-340). Newcastle: Cambridge Scholars.

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the*

ACM 26(11), 832-843.

Allen, J. F. & Ferguson, G. (1994). Actions and events in interval temporal logic. *Journal of Logic and Computation 4*(5), 531-579.

ASD Simplified Technical English Specification ASD-STE 100 TM: International specification for the preparation of maintenance documentation in a controlled language. Issue 6. January 2013. Brussels: ASD.

Ast, M. (2012). An ontology for aircraft design. (Diploma Thesis). <https://github.com/astbhltum/Aircraft-Ontology> [14/10/2016], under the Eclipse Public License.

Ast, M., Glas, M., Roehm, T., & Luftfahrt V, B. (2013). Creating an ontology for aircraft design. *DocumentID: 301356,* Deutsche Gesellschaft für Luft-und Raumfahrt-Lilienthal-Oberth eV.

Brusilovsky, P., & Cooper, D. W. (2002, January). Domain, task, and user models for an adaptive hypermedia performance support system. In *Proceedings of the 7th international conference on Intelligent user interfaces* (pp. 23-30). ACM.

Carrión-Delgado, M. G. & Felices-Lago, Á. (2014). La jerarquización de las entidades en la ontología satélite del crimen organizado y el terrorismo en FunGramKB. In Ch. Vargas Sierra (Ed.) *TIC, Trabajo colaborativo e interacción en terminología y traducción* (pp. 591-607). Granada: Comares.

Durak, U., Schmidt, A., & Pawletta, T. (2014). Ontology for objective flight simulator fidelity evaluation. *SNE Simulation Notes Europe*, *24*(2), 69-78.

Felices-Lago, Á. (2015). Foundational considerations for the development of the *Globalcrimeterm* subontology: A research project based on FunGramKB. *Onomázein*, *31*, 127-144.

Felices-Lago, Á. (2016). The process of constructing ontological meaning based on criminal law verbs. *Círculo de Lingüística Aplicada a la Comunicación*, *65*, 109-148.

Felices-Lago, Á. & Ureña-Gómez-Moreno, P. (2012). Fundamentos metodológicos de la creación subontológica en FunGramKB. *Onomázein, 26*(2), 49-67.

Felices-Lago, Á. & Ureña-Gómez-Moreno, P. (2014). FunGramKB Term Extractor: a key instrument for building a satellite ontology based on a specialized corpus. In B. Nolan & C. Periñán-Pascual (Eds.), *Language processing and grammars: The role of functionally oriented computational models* (pp. 251-269). Amsterdam: John Benjamins.

Gangemi, A., Sagri, M. and Tiscornia, D. (2005). A constructive framework for legal ontologies. In V.R. Benjamins et al. (Eds.). *Law and the Semantic Web* (pp. 97-124). Berlin: Springer.

Garst Jr, W. J. (2009). Mind games: the ontology of aviation safety and its consequences (Doctoral dissertation, Virginia Tech).

Hongjun, Q., Hua, T., Xiaobing, G., & Xue, Y. (2007). Ontology for aircraft assembly process planning knowledge. *Journal-Northwestern Polytechnical University*, *25*(1), 32.

Keller, R. M. (2016, September). Ontologies for aviation data management. In *Digital Avionics*

Systems Conference (DASC), 2016 IEEE/AIAA 35th (pp. 1-9). IEEE.

Kuhn, T. (2014). A survey and classification of controlled natural languages. *Computational Linguistics*, *40*(1), 121-170.

Kumar, B., & Marshall, D. M. (2005). An illustrated dictionary of aviation. New York: McGraw-Hill.

Lenci, A. (Ed.) (2008). From context to meaning: distributional models of the lexicon in linguistics and cognitive science. *Italian Journal of Linguistics*, *20*(1).

Liu, J., Tang, S. Y., Li, Q., & Wang, W. P. (2009, September). An ontology for aircraft route planning. In *Advances in System Simulation, 2009. SIMUL'09. First International Conference on* (pp. 68-72). IEEE.

Møller, M.H. & Christoffersen, E. (2006). Building a controlled language lexicon for Danish. *LSP and Professional Communication*, 6(1), 26-37.

Niaraki, A. S., & Kim, K. (2009). Ontology based personalized route planning system using a multi-criteria decision making approach. *Expert Systems with Applications*, *36*(2), 2250-2259.

Niles, I., & Pease, A. (2001). Towards a standard upper ontology. In Proceedings of the International Conference on Formal Ontology in Information Systems-Volume 2001 (pp. 2-9). ACM.

Periñán-Pascual, C. (2015). The underpinnings of a composite measure for automatic term extraction: the case of SRC. *Terminology, 21*(2), 151-179.

Periñán-Pascual, C. (2017). Bridging the gap within text-data analytics: a computer environment for data analysis in linguistic research. *Revista de Lenguas para Fines Específico, 23*(2), 111-132.

Periñán-Pascual, C. & Arcas-Túnez, F. (2004). Meaning postulates in a lexico-conceptual knowledge base. *15th International Workshop on Databases and Expert Systems Applications* (pp. 38-42). IEEE, Los Alamitos (California),

Periñán-Pascual, C. & Arcas-Túnez, F. (2007). Cognitive modules of an NLP knowledge base for language understanding, *Procesamiento del Lenguaje Natural, 39*, 197-204.

Periñán-Pascual, C. & Arcas-Túnez, F. (2010a). Ontological commitments in FunGramKB. *Procesamiento del Lenguaje Natural*, *44*, 27-34.

Periñán-Pascual, C. & Arcas-Túnez, F. (2010b). The architecture of FunGramKB. In *Proceedings of the 7th International Conference on Language Resources and Evaluation* (pp. 2667-2674). European Language Resources Association.

Periñán-Pascual, C. & Arcas-Túnez, F. (2014). La ingeniería del conocimiento en el dominio legal: La construcción de una ontología satélite en FunGramKB. *Revista Signos. Estudios de Lingüística*, *47*(84), 113-139.

Periñán-Pascual, C. & Mairal-Usón, R. (2009). Bringing Role and Reference Grammar to natural language understanding. *Procesamiento del Lenguaje Natural*, *43*, 265-273.

Periñán-Pascual, C. & Mairal-Usón, R. (2010). La Gramática de COREL: un lenguaje de representación conceptual. *Onomázein*. *21*(1), 11-45.

Periñán-Pascual, C. & Mestre-Mestre, E. M. (2015). DEXTER: automatic extraction of domain-specific glossaries for language teaching. In *Proceedings of the VII Congreso Internacional de Lingüística de Corpus. Procedia - Social and Behavioral Sciences.* 198, 377-385.

Reiss, M., Moal, M., Barnard, Y., Ramu, J. Ph., & Froger, A. (2006). Using ontologies to conceptualize the aeronautical domain. In *Proceedings of the International Conference on Human-Computer Interaction in Aeronautics (HCI-AERO 2006)*, Seattle, WA, Toulouse: Cépaduès-Editions, 56-63.

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: principles and methods. *Data & knowledge engineering*, *25*(1-2), 161-197.

Studer, R., Grimm, S., & Abecker, A. (2007). Semantic web services. Concepts, technologies and applications. Berlin: Springer.

Travers, R. W. (2000). Pilot-centered phase of flight standardization. In *Proceedings of the International Conference on Human-Computer Interaction in Aeronautics* (pp. 51-56). Toulouse: Cépaduès-Edition.

van Putten, B-J, Wolfe, S & Dignum, V. (2008). An ontology for traffic flow management. In *Proceedings of the 8th Aviation Technology, Integration, and Operations Conference*, AIAA. Anchorage, September 2008.

Van Valin, R. D. Jr. (2005). *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.

Van Valin, R.D. Jr. & LaPolla, R. (1997). *Syntax, Structure, Meaning and Function*. Cambridge: Cambridge University Press.

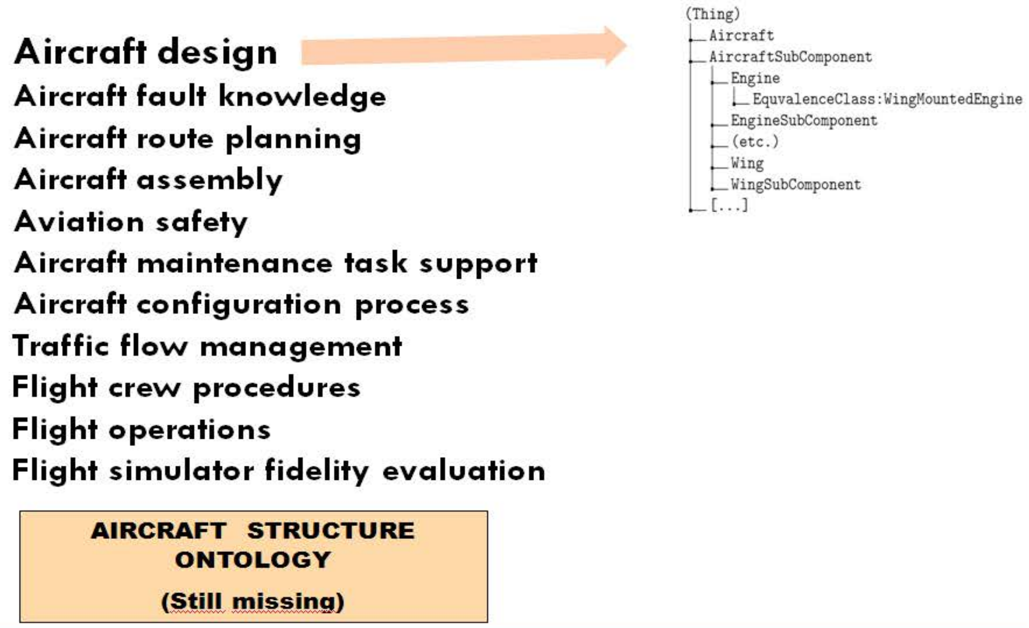
Wojcik, R., Holmback, H. & Hoard, J. (1998). Boeing technical English: An extension of AECMA SE beyond the aircraft maintenance domain. In *Proceedings of the Second International Workshop on Controlled Language Applications (CLAW-98)*, Pittsburgh, PA., 114-123.

Wu, Y., Ebrahimipour, V., & Yacout, S. (2014). Ontology-based modeling of aircraft to support maintenance management system. *IIE Annual Conference. Proceedings*. Institute of Industrial Engineers-Publisher, 1159-1168.

Yi, J., Ji, B., Chen, C., & Tian, X. (2009, March). Employing ontology to build the engine fault diagnosis expert system. *Computer Science and Information Engineering, 2009 WRI World Congress on* (Vol. 2, pp. 631-635). IEEE.

Zhou, Y., Zeng, Z. Y., Tian, B., Jia, Z. Y., & Guo, X. (2012). Ontology modeling of aircraft fault knowledge. In *Applied Mechanics and Materials* (Vol. 236, pp. 350-355). Zurich: Trans Tech Publications.

Appendix 1: Schema of existing ontologies in aeronautics



Appendix 2: Specialised sources for the construction of the Aircraft Structure Ontology

ATA [Air Transport Association]100 CHAPTER: Specifiction for manufacturers maintenance data.

HANDBOOKS (in paper and in electronic format)

- Reithmaier, L., & Sterkenburg, R. (2013). *Standard aircraft handbook for mechanics and technicians*. McGraw Hill Professional.
- Kroes, M., Watkins, W., Delp, F., & Sterkenburg, R. (2013). *Aircraft maintenance and repair*. McGraw Hill Professional.
- *Aviation Maintenance Technician Handbook* (U.S. Department of Transportation, Federal Aviation Administration, 2012)
    - General
    - Airframe, Volume 1 & 2
    - Powerplant, Volume 1 & 2

DICTIONARIES (in paper and in electronic format)

- Crane, D. (2012). *Dictionary of aeronautical terms*. Aviation Supplies & Academics, Incorporated. **(11,000 terms and nearly 500 illustrations)**
- Kumar, B., & Marshall, D. M. (2005). *An illustrated dictionary of aviation*. New York: McGraw-Hill. **(7,400 terms and 2,400 illustrations)**
- Crocker, D. (2010). *Dictionary of aviation.* London: A. & C Black. **(5,500 terms)**
- *Pilot/Controller Glossary* (2015). Federal Aviation Administration**.**

Appendix 3: Characteristics of the *Airbus corpus* and the automatic extraction with DEXTER.

AIRBUS CORPUS:

- Courtesy of *Airbus* in Seville
- 2,480 files / 6.697.387 bytes (xml format)
- 687,345 tokens
- Language: Written in Simplified Technical English (ASD-STE)
- Mode: written
- Type: synchronic, closed.
- Characteristics: Not tagged. Collection of raw texts for terminological extraction
- Domain: Aircraft maintenance.
- Subdomain: Aeronautical English **–** aircraft maintenance
- Instructions.
    - How to use…
    - Safety procedures
- Descriptions.
    - Elements
    - Technical data
    - System
- Warning notice

AUTOMATIC EXTRACTION:

- Tool: DEXTER **(D**iscovering and **EX**tracting **TER**minology**)** in the FunGramKB Suite

- Hybrid linguistic-statistical approach to terminology

(SRC: Salience, Relevance and Cohesion)

- Stop list for:
  - Functional words
  - Non-alphabetic characters
- List of n-gram candidates: 6,966
  - Unigrams: 1,887
  - Bigrams: 2,940
  - Trigrams: 2,139