Ana Ostroški Anić/Ivana Brač

AirFrame

Mapping the field of aviation through semantic frames

Abstract The paper presents the process of developing the AirFrame database, a specialized lexical resource in which aviation terminology is defined in the form of semantic frames, following the methodology of the Berkeley FrameNet (FN). First, the structure of the database is presented, and then the methodology applied in developing and populating the database is described. The link between specialized aviation frames and general language semantic frames, of which frames defining entities, processes, attributes and events are particularly relevant, is discussed on the example of the semantic frame of Flight and its related frames. The paper ends with discussing possibilities of using AirFrame as a model for further developing resources in which general and specialized knowledge are linked.

Keywords Terminology; aviation terminology; semantic frames; specialized knowledge; specialized lexicography

1. Introduction

Aviation is a professional domain in which language plays a crucial role in ensuring regular daily operations and safe communication. There is hardly any other professional environment in which the communicative setting is as described and prescribed as the one of the international aviation community. However, aviation language consists not only of aviation terminology and radiotelephony phraseology in English, but it also includes a certain level of general language vocabulary that is necessary for effective professional communication. Defining the lexical component of aviation language is therefore a challenging task because of all the different areas of expertise encompassed, but also because not all aviation subdomains are equally relevant for effective communication (Bratanić/Ostroški Anić 2010). In a similar manner, it could be claimed that any domain of specialized knowledge is characterized by different types of categories that are often intertwined with categories of general knowledge and human experience used in a specialized context.

The need for linking general and specialized knowledge has been well addressed in developing specialized resources based on the theory of Frame Semantics (Faber et al. 2011; L'Homme/Subirats/Robichaud 2016; L'Homme/Robichaud/Subirats 2020; Pilitsidou/Giouli 2020) or on its terminological application in the form of Frame-Based Terminology (Faber et al. 2011; Faber/Buendía Castro 2014; Faber 2015). Unlike traditionally organized specialized dictionaries or terminological databases that define professional knowledge through hierarchically organized data categories, interpreting specialized categories as dynamic structures calls for providing a more dynamic approach to the processing and presentation of terminology (Faber 2015). Although semantic relations between frames and their elements are also largely subject to a hierarchical structure, they nevertheless allow for including different chronological and associative relations (Ruppenhofer et al. 2016).

AirFrame is a specialized lexical resource in which the domain of aviation is defined in terms of semantic frames, i.e. specialized aviation frames are linked to general language semantic frames, of which top level frames defining various entities, processes, attributes

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

and events are most relevant (Brač/Ostroški Anić 2019). The AirFrame database was designed with a view of defining a specialized domain, but at the same time linking it to more general linguistic information, paving the way for concept modelling that enables linking resources of a different origin and a different purpose. This paper presents the structure of the database and the methodology applied in developing it, with a particular focus on links between general and specialized semantic frames defined in it. After the introduction, a brief overview of the application of FrameNet's methodology to defining specialized knowledge is given. The structure of semantic frames and their elements is presented in the third section, while the methodology of frames identification and description is given in the fourth section of the paper. Discussion is based on the example of the sematic frame of Flight and its related frames. The paper ends with suggestions for possible applications of the database data.

2. Application of FrameNet to specialized domains

There have been many applications of Frame Semantics to lexicography since Fillmore first defined the *frame* as "a system of categories structured in accordance with some motivating context" (Fillmore 1982), identifying experience as the context necessary for successful categorization. In the fields of terminology, specialized lexicography and specialized knowedge representation, a number of applications of Frame Semantics to specialized resources have been developed in order to enable a more accurate linguistic and computational representation of the conceptual level of specialized knowledge.

Kiktionary (Schmidt 2007) is an English-German-French database of football vocabulary extracted from the corpus of reports from football matches, in which FrameNet's methodology is combined with the Wordnet's principle of organizing synonyms and homonyms in synsets. BioFrameNet (Dolby/Ellsworth/Scheffczyk 2006) is another database applying the FrameNet's methodology in processing texts in the field of biomedical sciences and molecular biology. The domain of law is represented in, among others, the Italian resource combining Frame Semantics and Van Kraling's approach to the representation of law (Venturi et al. 2009) and JuriDiCo (Pimentel 2015), a database of Portuguese and English legal terminology modelled in accordance with the methodologies applied in FrameNet and the Canadian database DiCoInfo (L'Homme 2012).

Frame-Based Terminology emerged as a theoretical approach in terminology studies, applying the principles of Frame Semantics to terminology work by developing a model of dynamic description of categories of specialized knowledge, (Faber Benítez/Márquez Linares/Vega Expósito 2005; Faber 2015) where focus is put on the description of a prototypical event in a specialized domain, e. g. the environmental event in the domain of environment (Faber/Buendía Castro 2014). The prototypical event serves as a general frame for the organization of more specific concepts, and it encompasses macro-categories as "concept roles characteristic of this specialized domain" (Araúz/Reimerink/Faber 2009). In Puertoterm, a knowledge base on environmental engineering, an event includes a natural agent and a human agent, a process with the subcategories of a natural and an artificial process and a construct, and a patient/result. The human agent can use an instrument to create an artificial process or constructs (Araúz/Reimerink/Faber 2009). In Ecolexicon, an environmental knowledge base, the same dynamic conceptualization of the event structure is applied on all levels, which means that terminographic definitions can be also considered as "mini-knowledge representations or frames" (Araúz/Reimerink/Faber 2009, p. 51).

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

Termframe, a terminology knowledge base for the field of karstology, is another resource developed on the principles of Frame-Based Terminology, containing terms and their definitions in English, Slovene and Croatian. In Termframe, a definition template is defined for each concept category in the domain model (Vintar/Stepišnik 2021), following the frame-based approach that views a definition as a small frame of knowledge. What makes Termframe particulary relevant for our work is the development of first frame-based definition templates for Croatian, which could be adapted for terminology work in another specialized domain.

The work done by L'Homme and colleagues in developing domain-specific resources as applications of the Frame Semantics methodology and theoretical principles (L'Homme/Robichaud/Rüggeberg 2014; L' Homme/Robichaud/Subirats 2020) was of greatest influence in devising the AirFrame's methodology and data categories structure. Their model of linking general language semantic frames with domain-specific frames relies on using a set of 15 semantic roles, general enough in their description to be applicable to a number of terms in more semantic frames (Pimentel/L'Homme/Laneville 2012), and it served as the starting point for developing a methodology of linking aviation related semantic frames to their general top-level instances. The next section describes the structure of the AirFrame database in more detail.

3. The structure of AirFrame

AirFrame is the first specialized frame-based lexical resource in the Croatian language, consisting of aviation related semantic frames and frame elements (FEs) with their accompanying definitions and examples, types of frame elements, lexical units and frame-to-frame relations.[1] Since the domain of aviation is typically characterized by numerous events, activities and well-controlled processes, FrameNet's methodology of describing knowledge categories in terms of hierarchically related events and activities they structure seemed particularly appropriate for the identification of this field.

In designing the data categories structure, we tried to adhere to the FrameNet's structure as close as possible, but at the same time to use the categories that reflect the basic principles of terminology work. Apart from administrative categories, each frame therefore consists of a frame definition, frame elements with their definitions, examples of sentences in which the frame specific frame elements are used, lexical units and relations to other frames. AirFrame is primarily a database of Croatian aviation terminology, so the definitions of frames and frame elements are written in Croatian, as are the types of frame elements and frame-to-frame relations. However, examples of frame elements and all lexical units are given for Croatian, English and French alike.

Following the FrameNet distinction, frame elements can be core, which are frame specific and inessential in the realization of a given frame, and non-core, i. e. elements that do not uniquely characterize the frame (Ruppenhofer et al. 2016), but further define it by placing it in a certain time and space. A further distinction between non-core and peripheral elements is not applied, nor are extra-thematic elements defined.

Another divergence from the FrameNet model is the omission of annotated examples for FEs lexical realization. Although the annotation of sentences as examples of lexical units in

[1] AirFrame is available at airframe.jezik.hr.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

linguistic context has been done, it is not yet integrated into the database. The category of frame element example therefore does not illustrate valency patterns for particular lexical units that can appear in syntactic position of a defined element, as it does in FrameNet, but it resembles the category of context as a typical element of a traditional terminology database. A defined FE is still placed within an actual linguistic context, but the difference between a terminological context and an example of an FE in AirFrame lies in the fact that the example can contain any of the lexical units – i. e. all synonyms and variants of a term – of the same meaning in the frame. All instances of examples are corpus examples, and can be given for Croatian, English and French. There is a difference in the approach to example illustration between aviation specific frames and general language frames, which is discussed in the next section.

Since no ontology has yet been developed as a formal model of conceptual representation in AirFrame, the category of the frame element type was introduced to serve as the connection between the semantic and ontological levels, i. e. as an implicit top level ontology. A set of 17 frame element types was composed on the basis of FrameNet's semantic types, EuroWordNet's top level entities (Vossen 2002) and the semantic roles defined in the LIRICS project (Petukhova/Bunt 2008). Unlike the ontological semantic types for the frame elements in FrameNet that categorize the sort of filler expected in the element (Ruppenhofer et al. 2016, p. 86), the FE type attributed to an FE in AirFrame classifies the kind of role the FE is. FE types in that sense serve as the superordinate macro roles for the frame specific semantic roles or frame elements, and are used to group elements in conceptually connected groups. E. g., the frame `Flight` has several frame elements referring to space and time: AIRSPACE and AERODROME are core elements bearing the FE type of location, as well as the non-core element FLIGHT_ROUTE.[2] The core element FLIGHT_TIME and non-core elements FLIGHT_DURATION, FREQUENCY and TIME_SPAN are marked with the FE type time.

Semantic frames are invoked by lexical units, which in AirFrame can be words of general language or aviation terms, and all instantiate FEs in appropriate frames. They are entered into the database separately from the frames, so they can be attributed to more than one frame. Lexical units are defined for Croatian, English and French, but only grammatical information for each is entered. When compared to the terminological information in a traditional termbase, lexical units would correspond to terms, while frame elements would roughly be on the level of concepts. However, a one-to-one relation cannot be established since FEs have the role of frame specific semantic roles.

Finally, all frames are linked by 13 FrameNet's frame-to-frame relations. Some relations are more relevant for the identification of conceptual relations between domain categories, such as *has subframe* and *subframe_of*, corresponding to partitive terminological relations, as well as *inherits* and *is_inherited_by*, which could be compared to generic relations, e. g. the type_of relation. Other relevant frame relations are shown in section 5.

4. Methodology for defining specialized semantic frames in AirFrame

The process of describing aviation semantic frames can be broadly divided into two large segments: identifying frames and defining them. Before starting any terminology work,

[2] The names of frames are written in fixed-width font, `Courier New`, while the names of frame elements are written in SMALL CAPS.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

regardless of the type of resource envisaged as the final product, one needs to first have a general overview of the domain whose terminology is being processed. The development of AirFrame was made easier in that sense because it continued from the work previously done on the Croatian aviation terminology (Ostroški Anić 2020).

The field of aviation was first broadly divided into large categories according to the basic processes and entities included and connected to the central event of the domain, i. e. the flight. Most aviation training material, e. g. handbooks, manuals and guidance material, are organized in a similar fashion, i. e. in a way which follows an aircraft from pre-flight to post-flight activities. This top-down approach is supported by a corpus-based analysis, for which two corpora had been compiled.

First, a parallel English-Croatian corpus was compiled using documents from the Directory of legal acts of the European Union, chapter Transport policy, subchapter Air transport in English and Croatian. Out of 220 documents from the Air transport subchapter, 178 legal acts were taken having both (English and Croatian) language versions. The texts were downloaded from the EUR-Lex database, and entered in the Sketch Engine's corpus compilation module (Kilgarriff et al. 2014). The English part of the corpus consists of a little over 950 000 words, while the Croatian part consists of 855 000 words. A monolingual corpus of aviation related texts in Croatian was compiled, too, for which the available textbooks, manuals, reports, scientific papers and dissertations, as well as student diploma papers and MA theses in Croatian were used. The aviation corpus in Croatian consists of 2,210,000 words.[3]

The parallel corpus was first used for term extraction and validation of term candidates' lists (Ostroški Anić/Lončar/Pavić 2019). An automatic term extraction was conducted for each language with the option of extracting a list of 1000 single-word and multi-word keywords. The EUR-Lex English 2/2016 corpus was used as a reference corpus for extracting the English single-word term candidates, while the English Web 2013 was used as a reference corpus for extracting the multi-word term candidates. Similar options were possible for term extraction in Croatian. Manual analysis and term verification of both English and Croatian candidate lists of extracted terms was then conducted.

After acquiring a list of terms or lexical units from the parallel corpus, a similar process was done for the Croatian aviation corpus. The list of terms extracted from the monolingual corpus served for the validation of terms from the parallel corpus. Given the nature of the legal discourse of the Eur-Lex documents in the parallel corpus, few definitions of aviation concepts were used in the description of semantic frames and their elements. The Croatian aviation corpus was used for this instead.

The semantic frames were then identified by grouping the extracted lexical units according to the aviation concepts they denote. In FrameNet, the basic criterion for delimiting one frame from another is that all lexical units should "evoke the same type of event and share the same inventory and configuration of FEs" (Ruppenhofer/Boas/Baker 2014). Therefore, lexical units of the same semantic type, and appearing as arguments of the same verbs should be placed in the same frame. Determining the scope of the semantic frame, however, largely depends on the granularity of the conceptual description, which is, on the other hand, conditioned by the potential use of the database.

Adding general language frames to AirFrame is done for two reasons: the first is the top-level categorization that provides a continuity in relations between more general frames and

[3] Both corpora are available to Sketch Engine users by contacting the authors.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

specialized frames as their instantiations. The second reason is laying down the foundations for a future FrameNet of the Croatian language, which could use the existing top-level frames, and enrich them if necessary.

The frames of general language are of the same structure as the aviation frames, except that they are not identified in the same way, but are taken over from FrameNet and adapted for Croatian. If all the frame elements of a certain frame from the Berkeley FN are appropriate for the description of the Croatian counterpart, nothing is changed. If the Croatian syntactic description asks for an additional element or a change in the existing FEs structure, changes are introduced accordingly. The Croatian general language corpus hrWaC is used for the examples of FEs (Ljubešić/Klubička 2016). English and French examples of FEs and lexical units, otherwise added in aviation semantic frames, are not entered in the general language frames. Users can look for this information in the English and French FrameNet.

5. The semantic frame of Flight

Flight is without doubt the central event in the domain of aviation. There are several perspectives one can take in order to define its activities, processes and entities engaged. Flight can be defined as an instance of travel, in which case we are taking the view of passengers and defining it as the period of transport by aircraft from boarding the airplane to its disembarkation. Flight can also be defined by taking into account the cargo and baggage being transported, but in AirFrame it is described with regard to the use of the aircraft by authorized aviation personnel to conduct the activity of operating the aircraft for flying from one location to another.

As already said, every frame has a definition that contains all core FEs that are conceptually necessary to understand the frame. The definition of `Flight` in AirFrame is: The AIRCRAFT moves from a specific AERODROME to a specific destination AERODROME through AIRSPACE for a specified DURATION. The definition tells us that the core frame elements are: AIRSPACE, AERODROME, AIRCRAFT, and DURATION, while non-core elements include, among others: PILOT, ALTITUDE, FLIGHT_ROUTE, FLIGHT_PATH, FLIGHT_CONDITIONS, FLIGHT_CREW, MANNER, FLIGHT_SPEED, FREQUENCY, TIME_SPAN, etc. FEs are inferred from corpus examples, and are defined in relation to a specific frame, e. g. AIRCRAFT is defined as 'a device that can maintain istelf and move in atmosphere', FLIGHT_SPEED as 'the speed of the aircraft in relation to the Earth's surface', etc. Below every definition of a frame element, there are examples in Croatian, English, and French to show its use in context, as shown in Figure 1.

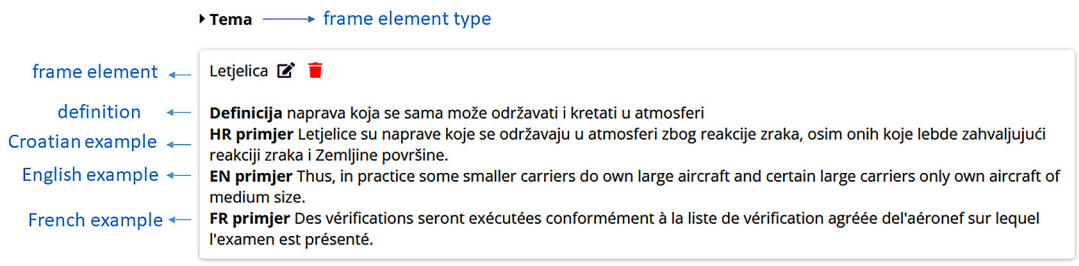

Fig. 1: Frame element *Letjelica* (eng. Aircraft)

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

Different frame elements are grouped under frame element types, understood here as macro roles connecting semantic and ontological levels of information. In `Flight`, the frame element type Location encompasses AIRSPACE and AERODROME as core elements of `Flight`, and FLIGHT_ROUTE as a non-core element. MANNER and SPEED fall under the FE type Manner, and FREQUENCY and TIME_SPAN under Time. A general FE type of Theme is attributed to the core-element AIRCRAFT when used in that semantic role, while the PILOT is a specification of the FE type of Agent.

Although semantic annotation of corpus examples showing the argument structure of FEs is going to be added to the database in the next phase of the development, a significant number of sentences has already been annotated in a separate database. Examples (1) to (6) show the annotation of sentences with the target units FLIGHT and FLY. Frame elements are marked with a subscript, while frame element types are written in superscript.

(1) Avioni predviđeni za let na visinama iznad 25,000 stopa moraju biti opskrbljeni jedinicama za raspodjelu kisika.

$^{\text{THEME}}[_{\text{AIRCRAFT}}$ Airplanes] intended for FLIGHT $^{\text{LOCATION}}[_{\text{ALTITUDE}}$ at altitudes above 25,000 feet] must be equipped with oxygen distribution units.

(2) Ako pilot zrakoplova radi sigurnosnih ili hitnih letačko-operativnih razloga smatra da nastavak leta prema izvornom aerodromu odredišta nije preporučljiv, može preusmjeriti let na drugi aerodrom, kojega smatra prikladnim.

If $^{\text{AGENT}}[_{\text{PILOT}}$ the pilot of the aircraft], $^{\text{CAUSE}}[_{\text{CAUSE}}$ for safety or emergency flight-operational reasons], considers that the continuation of the FLIGHT $^{\text{FINAL\_LOCATION}}[_{\text{AERODROME}}$ to the original destination aerodrome] is not recommended, $^{\text{AGENT}}[_{\text{PILOT}}$ he] may redirect the FLIGHT $^{\text{FINAL\_LOCATION}}[_{\text{AERODROME}}$ to another aerodrome] which he considers appropriate.

(3) Putanja dopušta da helikopter nastavi let od visine krstarenja do visine od 300 m (1000 ft) iznad helidroma.

$^{\text{THEME}}[_{\text{FLIGHT\_PATH}}$ The trajectory] allows $^{\text{THEME}}[_{\text{AIRCRAFT}}$ the helicopter] to continue its FLIGHT $^{\text{INITIAL\_LOCATION}}[_{\text{ALTITUDE}}$ from the height of the cruise] $^{\text{FINAL\_LOCATION}}[_{\text{ALTITUDE}}$ to an altitude of 300 m (1000 ft) above the heliport].

(4) Finnair je putnicima koji su u SAD letjeli preko Helsinkija jedno vrijeme poklanjao SharpWizard 8000.

$^{\text{AGENT}}[_{\text{AIRLINE}}$ Finnair] gave away $^{\text{THEME}}$[SharpWizard 8000] $^{\text{GOAL}}$ $[_{\text{PASSENGER}}$ to passengers] who FLEW $^{\text{FINAL\_LOCATION}}$[to the USA] $^{\text{LOCATION}}[_{\text{FLIGHT\_ROUTE}}$ via Helsinki] $^{\text{TIME}}[_{\text{TIME}}$ for a while].[4]

(5) Kandidat mora letjeti helikopterom najmanje 5 sati noću, od čega najmanje 3 sata s instruktorom, uključujući 1 sat rutnog navigacijskog letenja te 5 samostalnih polijetanja i slijetanja sa zaustavljanjem.

$^{\text{AGENT}}[_{\text{PILOT}}$ The candidate] must FLY $^{\text{THEME}}[_{\text{AIRCRAFT}}$ the helicopter] $^{\text{TIME}}[_{\text{DURATION}}$ for at least 5 hours at night], of which $^{\text{TIME}}[_{\text{DURATION}}$ at least 3 hours] $^{\text{AGENT}}[_{\text{PILOT}}$ with an instructor], including $^{\text{AMOUNT}}[_{\text{DURATION}}$ 1 hour] $[_{\text{FLIGHT}}$ of route navigational flight] and $^{\text{AMOUNT}}[_{\text{AMOUNT}}$ 5] $[_{\text{FLIGHT\_SEGMENT}}$ solo take-offs and landings with stopping].

(6) Prvi američki mlazni zrakoplov, Boeing B-707, proizveden je 1958. godine i letio je uglavnom preko Sjevernog Atlantika.

$^{\text{THEME}}[_{\text{AIRCRAFT}}$ The first American jet, the Boeing B-707], was built $^{\text{TIME}}[_{\text{TIME}}$ in 1958] and FLEW mostly $^{\text{LOCATION}}[_{\text{FLIGHT\_ROUTE}}$ across the North Atlantic].

[4] The frame element type Goal also covers the semantic roles Recipient and Beneficiary, which are attributed to animate participants of an event.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

There are many lexical units evoking the frame `Flight`, as can be seen in Figure 2. There is no terminological preference given to certain lexical units over others, the way terms are usually classified according to normative preference in a traditional terminological resource. All synonyms and variants of a given term can therefore be added as lexical units because they all do evoke a semantic frame for which the term is relevant. However, spelling and formal variants are usually omitted.

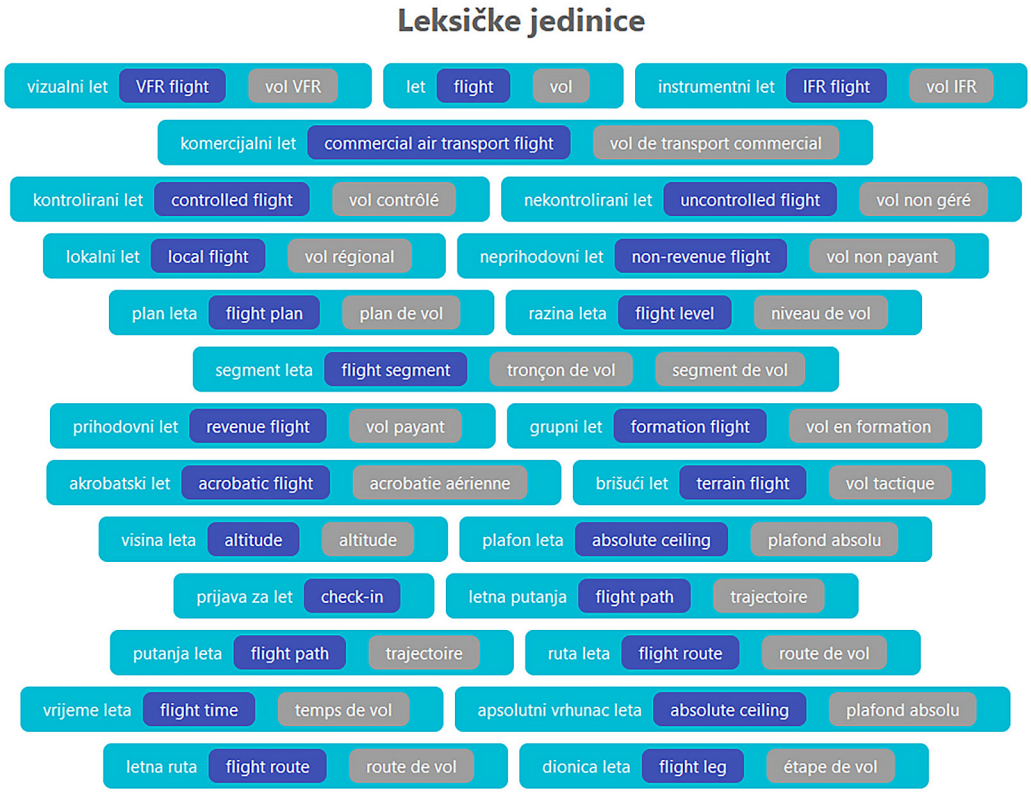

Fig. 2: Croatian, English and French lexical units of the frame `Let` (eng. `Flight`) represented in the search engine airframe.jezik.hr

Since one of the aims of AirFrame is to present the structure of the field of aviation as comprehensively as possible, frame-to-frame relations have been included in the frame description. `Flight` inherits all FEs from top-level frames `Event` and `Motion` as its parent frames, with more specific FEs added. Being a complex frame that an event is, `Flight` can be further divided into subframes, i. e. into six subframes that correlate to the phases of flight: `Take-off`, `Landing`, `Climb`, `Cruising`, `Descent`, `Approach`. Finally, `Flight` uses certain elements of the frames `Airspace` and `Airport`, while the frame `Flight_travel` is closely related to it and therefore linked by the relation *see_also*. Frame-to-frame relations used to connect `Flight` to related frames in the AirFrame database are shown in Figure 3.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

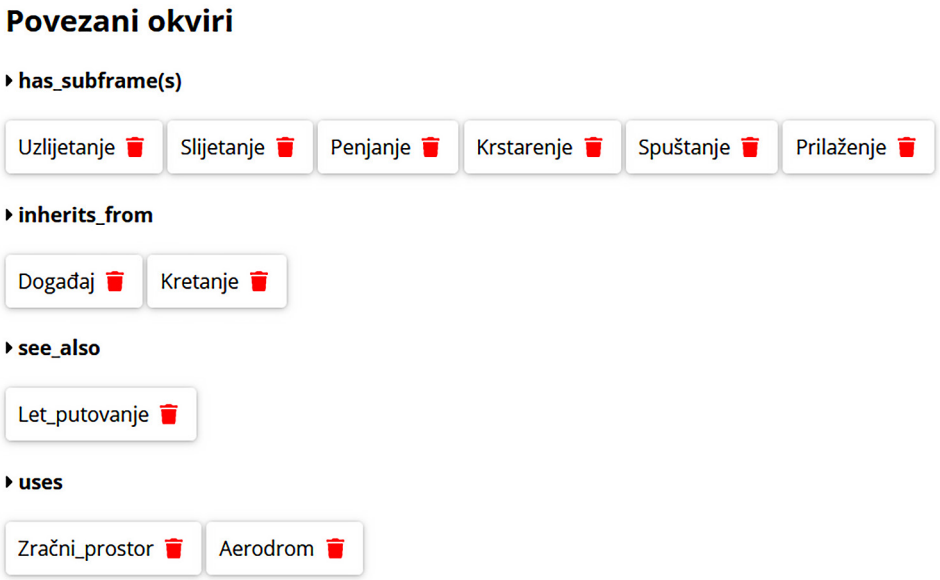

Fig. 3: Frames related to `Flight`

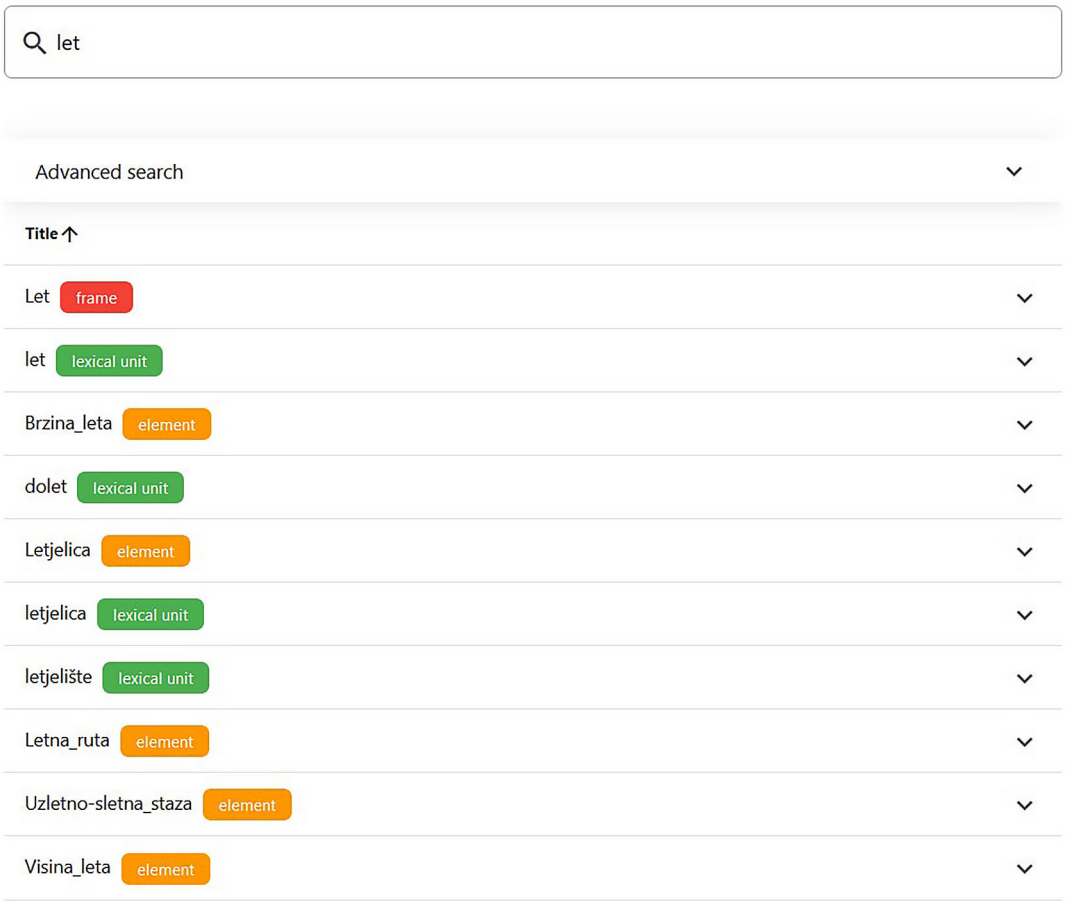

Fig. 4: Search results for *let* 'flight' in the AirFrame search engine

The analysis of such a complex category that combines different spatial and temporal relations is a good example of the possibility of connecting extra-linguistic and linguistic

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

knowledge in the form of semantic frames, which can be used in a number of different applications. When used as a teaching and reference resource, AirFrame's offers a more outlined presentation of the conceptual level of specialized knowledge than users find in traditional specialized resources. As can be seen in Figure 4, users can search for any term related to the frames, frame elements and lexical units defined, while the display of results visually keeps the distinction between different categories in the database, as well as between lexical units in different languages.

6. Concluding remarks

Specialized knowledge defined and presented in a domain-specific lexical resource such as the one presented here provides ample opportunities for further use. However, the envisaged users of AirFrame, as well as users of all similar specialized resources, are not typical terminology users relying on intuitive data categories presentation that enables them to quickly find answers to various terminological issues. Domain lexical resources are often more used by machines than humans, perhaps because they provide a thorough representation of a complex network of semantic relations, which lies underneath all linguistic structures. Their potential is nevertheless more and more recognized in developing LSP courses and teaching material. Although students of aeronautics, aviation engineering, air traffic and related studies, as well as translators working in these fields, are without doubt primary intended users of AirFrame, the database is being developed as a model of specialized knowledge description that can be used in the description of other, related specialized domains.

AirFrame follows the footsteps of several similar specialized resources that apply Frame Semantics and versions of the FrameNet's methodology in particular. What sets it apart is the introduction of the level of general semantic roles that can be used for linking different resources that use semantic role labelling, whether those are general language or domain-specific resources. Using the frame specific semantic roles gives us a detailed insight into the structure of categories, and allows for mutual relations to be established between different semantic frames.

Grouping frame specific elements into higher level semantic roles, on the other hand, gives us a list of all elements of the same type that appear in different syntactic functions. For example, the lexical units *airplane*, *helicopter* or *jet*, which in the examples presented in the previous section are defined as belonging to the FE AIRCRAFT, and to which the FE type Theme is assigned, may appear in the subject's position in cases where no agent is expressed. In other words, it is not expressed that someone flies the aircraft. In sentences where an agent is expressed, i. e. in sentences where the FE PILOT is realized by the lexical units *pilot* and *candidate*, the agent appears in the object position. Based on the comparison of all FE within the same FE type, we can thus make conclusions on different syntactic patterns in which they appear, and whether a change in the lexical unit used for a particular FE changes the intended meaning of the whole utterance.

A combination of a fine-grained semantic description with an enabled level of generalization might pave the road to the development of a robust conceptual model that preserves terminological information in the integration of terminological resources into larger linguistic networks.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

References

Araúz, P. L./Reimerink, A./Faber, P. (2009): Puertoterm & Marcocosta: a frame-based knowledge base for the environmental domain. In: Proceedings of the XVIII FIT World Congress. Shanghai, p. 24.

Brač, I./Ostroški Anić, A. (2019): From concept definitions to semantic role labeling in specialized knowledge resources. In: Gürlek, M./Naim Çiçekler, A./Taşdemir, Y. (eds.): Proceedings of the 13th International Conference of the Asian Association for Lexicography, pp. 604–611.

Bratanić, M./Ostroški Anić, A. (2010): Compiling lexical information for an aviation English dictionary. In: Reconceptualizing LSP. Online proceedings of the XVII European LSP Symposium 2009.

Dolby, A./Ellsworth, M./Scheffczyk, J. (2006): BioFrameNet: a domain-specific FrameNet extension with links to biomedical ontologies. In: Bodenreider, O. (ed.): Proceedings of KR-MED. Formal Biomedical Knowledge Representation. Baltimore, pp. 87–94.

Faber Benítez, P./Márquez Linares, C./Vega Expósito, M. (2005): Framing terminology: A process-oriented approach. In: Meta: Journal des traducteurs 50 (4). https://www.erudit.org/fr/revues/meta/2005-v50-n4-meta1024/019916ar/ (last access: 29-06-2022).

Faber, P. (2015): Frames as a framework for terminology. In: Kockaert, H. J./Steurs, F. (eds.): Handbook of terminology. Amsterdam, pp. 14–33. http://doi.org/10.1075/hot.1.02fra1.

Faber, P. et al. (2011): Linking specialized knowledge and general knowledge in EcoLexicon. In: Actes de la conférence Terminologie & Ontologie: Théories et Applications (TOTh) 2011. Annecy, pp. 47–61.

Faber, P./Buendía Castro, M. (2014): EcoLexicon. In: Abel, A./Vettori, C./Ralli, N. (eds.): Proceedings of the 16th EURALEX International Congress. Bolzano, pp. 601–607. https://euralex.org/publications/ecolexicon/.

Fillmore, C. J. (1982): Frame semantics. In: Linguistics in the morning calm. Selected papers from SICOL-1981. Seoul, Korea, pp. 111–137.

Kilgarriff, A. et al. (2014): The Sketch Engine: ten years on. In: Lexicography 1 (1), pp. 7–36. http://doi.org/10.1007/s40607-014-0009-9.

L 'Homme, M.-C./Robichaud, B./Subirats, C. (2020): Building multilingual specialized resources based on FrameNet: application to the field of the environment. In: Torrent, T. T./Baker, C. F./Czulo, O./Ohara, K./Petruck M. R. L. (eds.): International FrameNet Workshop 2020 Towards a Global, Multilingual FrameNet. Marseille, pp. 85–92.

L' Homme, M.-C./Subirats, C./Robichaud, B. (2016): A proposal for combining "general" and specialized frames. In: Proceedings of the 5th Workshop on Cognitive Aspects of the Lexicon (CogALex – V). Osaka, Japan, pp. 156–165. https://aclanthology.org/W16-5321.

L'Homme, M.-C. (2012): Adding syntactico-semantic information to specialized dictionaries: an application of the FrameNet methodology. In: Lexicographica 28, pp. 233–252. http://doi.org/10.1515/lexi.2012-0012.

L'Homme, M.-C./Robichaud, B./Rüggeberg, C. S. (2014): Discovering frames in specialized domains. In: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). Reykjavik, Iceland, pp. 1364–1371. http://www.lrec-conf.org/proceedings/lrec2014/pdf/455_Paper.pdf.

Ljubešić, N./Klubička, F. (2016): Croatian web corpus hrWaC 2.1. In: http://nlp.ffzg.hr/resources/corpora/hrwac/, https://www.clarin.si/repository/xmlui/handle/11356/1064.

Ostroški Anić, A. (2020): Hrvatsko zrakoplovno nazivlje od leksikografije do terminologije. In: Brač, I./Ostroški Anić, A. (eds.): Svijet od riječi. Terminološki i leksikografski ogledi. Zagreb, pp. 361–374.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.

Ostroški Anić, A./Lončar, M./Pavić, M. (2019): Extracting lexical units for identifying specialized semantic frames. In: Terminologija 26, pp. 73–87.

Petukhova, V./Bunt, H. (2008): LIRICS semantic role annotation: design and evaluation of a set of data categories. In: Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08). Marrakesh, pp. 39–45.

Pilitsidou, V./Giouli, V. (2020): Frame semantics in the specialized domain of finance: building a termbase to translation. In: Gavriilidou, Z./Mitsiaki, M./Fliatouras, A. (eds.): Euralex XIX. Lexicography for inclusion, pp. 263–271. https://euralex.org/publications/frame-semantics-in-the-specialized-domain-of-finance-building-a-termbase-to-translation/.

Pimentel, J. (2015): Using frame semantics to build a bilingual lexical resource on legal terminology. In: Kockaert, H. J./Steurs, F. (eds.): Handbook of Terminology. Amsterdam, pp. 427–450. http://doi.org/10.1075/hot.1.usi1.

Pimentel, J./L'Homme, M.-C./Laneville, M.-È. (2012): General and specialized lexical resources: a study on the potential of combining efforts to enrich formal lexicons. In: International Journal of Lexicography 25 (2), pp. 152–190. http://doi.org/10.1093/ijl/ecr025.

Ruppenhofer, J. et al. (2016): FrameNet II: extended theory and practice. https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf (last access: 07-03-2022).

Ruppenhofer, J./Boas, H. C./Baker, C. (2014): The FrameNet approach to relating syntax and semantics. In: Gouws, R. H./Heid, U./Schweickard, W./Wiegand, H. E. (eds.): Dictionaries. An international encyclopedia of lexicography. Berlin/New York, pp. 1320–1329.

Schmidt, T. (2007): The Kicktionary: a multilingual lexical resource of the language of football. In: Data Structures for Linguistic Resources and Applications: Proceedings of the Biennial GLDV Conference 2007. Tübingen, pp. 189–196.

Venturi, G. et al. (2009): Towards a FrameNet resource for the legal domain. In: Proceedings of the 3rd Workshop on Legal Ontologies and Artificial Intelligence Techniques: 2nd Workshop on Semantic Processing of Legal Text., p. 10. http://ceur-ws.org/Vol-465/paper8.pdf.

Vintar, Š./Stepišnik, U. (2021): TermFrame: a systematic approach to Karst Terminology. In: Dela (54), pp. 149–167. http://doi.org/10.4312/dela.54.149-167.

Vossen, P. (2002): EuroWordNet General Document. https://research.vu.nl/ws/portalfiles/portal/77020259/EWNGeneral.

Contact information

Ana Ostroški Anić
Institute of Croatian Language and Linguistics
aostrosk@ihjj.hr

Ivana Brač
Institute of Croatian Language and Linguistics
ibrac@ihjj.hr

Acknowledgements

This work has been fully supported by the Croatian Science Foundation under the project UIP-2017-05-7169.

This paper is part of the publication: Klosa-Kückelhaus, Annette/Engelberg, Stefan/Möhrs, Christine/Storjohann, Petra (eds.) (2022): Dictionaries and Society. Proceedings of the XX EURALEX International Congress. Mannheim: IDS-Verlag.