Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), pages 953–957
Marseille, 11–16 May 2020

An Annotated Corpus of Adjective-Adverb Interfaces in Romance Languages

Katharina Gerhalter[1], Gerlinde Schneider[2], Christopher Pollin[2], Martin Hummel[1]
[1]Department of Romance Studies, University of Graz, Austria
[1]Centre for Information Modelling, University of Graz, Austria
{katharina.gerhalter, gerlinde.schneider, christopher.pollin, martin.hummel}@uni-graz.at

Abstract

The final outcome of the project *Open Access Database: Adjective-Adverb Interfaces in Romance* is an annotated and lemmatised corpus of various linguistic phenomena related to Romance adjectives with adverbial functions. The data is published under open-access and aims to serve linguistic research based on transparent and accessible corpus-based data. The annotation model was developed to offer a cross-linguistic categorization model for the heterogeneous word-class "adverb", based on its diverse forms, functions and meanings. The project focuses on the interoperability and accessibility of data, with particular respect to reusability in the sense of the FAIR Data Principles. Topics presented by this paper include data compilation and creation, annotation in TEI/XML, data preservation and publication process by means of the GAMS repository and accessibility via a search interface. These aspects are tied together by semantic technologies, using an ontology-based approach.

Keywords: Corpus linguistics, Romance languages, Linguistic annotation, Semantic web, TEI/XML

1. Introduction

Romance languages have a large inventory of adverbs formed through different morphosyntactic word-formation processes based on adjectives. The basic process, inherited from Latin, consists of using the masculine singular form of the adjective for an adverbial function, that is, using an adjective in the syntactic position of an adverb. The following examples (meaning 'to see clear(ly)' and 'to fly high') show that the so-called adjective-adverbs are shared by several Romance languages (Hummel 2017):

1. Spanish: *ver claro*, *volar alto*
   Portuguese: *ver claro, voar alto*
   French: *voir clair, voler haut*
   Italian: *vedere chiaro, volare alto*
   Romanian: *a vedea chiar / clar, a zbura înalt*

On the other hand, several – but not all – Romance languages use the derivational suffix *-mente* as a specific adverbial word-formation process based on adjectives (Sp./Pt. *claramente*, Fr. *clairement,* It. *chiaramente* 'clearly'). Furthermore, the pattern "preposition + adjective" produced some prepositional phrases that were lexicalized as adverbial locutions, for example Sp. *de seguro* 'surely', Pt. *de novo* 'again', Fr. *pour de vrai* 'certainly', It. *sul serio* 'seriously' and Rom. *cu derept* 'fairly and equitably' (Hummel et al. 2019).

Therefore, adjective-adverbs in Romance languages are an interesting phenomenon in terms of interfaces with other adverbial formations based on the same lexical stem. Since the word-class "adverb" in general is highly heterogeneous, it subsumes various syntactic functions. On the one hand, adjective-adverbs are used as manner adverbs that modify the verb (as in examples 1, 'to see clear(ly)' or 'to fly high'), but also as focus-adverbs or specifiers (Fr. *juste une minute* 'just a minute') or as discourse markers (Sp. *Necesitará dinero, claro, como siempre* 'S/he needs money, sure / clearly, as always').

To study the mentioned phenomena, the research group "Adjective-Adverb Interfaces in Romance" has created relevant corpora in the course of various research projects. Within the project *Open Access Database* - "*Adjective-Adverb Interfaces in Romance*" (=AAIF)[1] these corpora of historical and present-day language examples have been compiled to one comprehensive resource for several Romance languages. The focus of this database lies on adjective-adverbs which are uniformly and comprehensively annotated and lemmatised. Prepositional phrases including an adjective and derived adverbs (e.g. *mente*-adverbs) were annotated in some corpora in order to be compared to adjective-adverbs.

The corpus can be used for cross-linguistic research on adjective-adverbs in several Romance languages. For example, the annotation of fine-grained morphosyntactic categories enables search queries for inflected adjective-adverbs. For example, the French adjective-adverbs *droit* 'straight' and *haut* 'high' show feminine inflection in the sentence *la flamme monte droite et haute* 'the flame rises straight and high'. This phenomenon is quite common in older texts and challenges linguistic definitions of the word-class adverb, which is supposed to be invariable. Therefore, morphology is not sufficient to distinguish between adjectives and adjective-adverbs, and inflection turns out to be one of the several interfaces between adjectives and adverbs. Hence, the definition of adverbs relies necessarily on their diverse syntactic functions, which were subsumed into several main functions (manner adverbs, speci-

[1] The project was funded by the pilot program *Open Research Data of the Austrian Science Fund* (FWF: ORD 66-VO). It is a cooperation between the *Department of Romance Studies* and the *Centre for Information Modelling – Austrian Centre for Digital Humanities*, both located at the University of Graz (Austria). Martin Hummel, the project leader, and Katharina Gerhalter take on the data collection and elaboration of linguistic categories for the annotation. Gerlinde Schneider and Christopher Pollin are in charge of the data modelling, the annotation tool and the processing and displaying of the data. https://adjective-adverb.uni-graz.at/en/projects/open-access-database-2017-2019.

fiers, discourse markers, etc.) in the annotation model.
In the following, we will present these corpora and their characteristics in detail, as well as describe the approach chosen for the linguistic annotation of the data. Furthermore, an important aspect of this project is that the research data is made accessible to the scholarly community on the web in a structured, comprehensible and open manner, thus following the FAIR data principles, Semantic Web technologies as well as the recommendations of CLARIN (De Jong et. al. 2018). The research data is made available through the digital repository GAMS, which is used for long-term archiving of research data in the humanities (Stigler & Steiner 2018).

2. Annotated Corpora

The AAIF-database currently contains 31.804 annotated examples for adjective-adverb interfaces in French, Spanish, Brazilian Portuguese, Varieties of Southern Italy (e.g. Neapolitan) and Romanian. It combines eight corpora which differ in temporal and regional coverage, source provenance, text type as well as historical or current language use. The selection of the examples is based on the scientific expertise of the domain experts with respect to their specific research questions and topics. Hence, the corpora are heterogeneous regarding the types of adverbs that are annotated. The project updates already analysed and partially tagged subcorpora and further includes newly tagged data by the project team and by cooperation partners (to be finished by the end of 2019). During the project, the annotation tool was updated in order to increase the annotation model with new categories.

2.1. Updating of previously collected corpora

The *Dictionnaire historique de l'adjectif-adverbe* (Hummel and Gazdik, in print) is based on language examples which have been available since 2005. The newly updated corpus (=Fr_A_DHAA) contains adjective-adverb examples from the 11th to the 20th century. The compilation was made on the basis of *Frantext Corpus* and *Corpus of the Dictionnaire du Moyen Français*. The annotation focuses on the verb phrase "*verb + adjective-adverb*" (e.g. *voir grand* 'to think big'), therefore, the corpus contains exclusively manner adverbs. The adjective-adverbs are lemmatized and tagged according to their morphological form and inflection, the number of syllables, coordination and modification of the adverb (e.g. *Le clocher sonnait plus clair* 'the bell (tower) sounded more clear(ly)'). In addition, the verbs of the phrases are also lemmatized and tagged (transitive, intransitive or reflexive, as well as coordination). The combination of both tagged verbs and tagged adjective-adverbs enables syntactic analysis (relative order of verb and adverb).

The *Corpus of French Adjective-Adverbs in informal texts from the Web* (=Fr_A_Web) adds present-day examples of colloquial French manner adverbs used in "*verb + adjective-adverb*" phrases, taken from web resources, such as blogs or forums. The examples were selected by a targeted web search, conducted during the works for the *Dictionnaire historique de l'adjectif-adverbe*. Tagging criteria apply the same classifications as the above mentioned corpus.

The *Corpus of Spanish Adjective-Adverbs in Diachrony* (= Sp_AP_SH3) used for the *Sintáxis Histórica III*-chapter "Adjetivos Adverbiales" (Hummel 2014) contains examples from the 13th to the 21st century. Examples were collected by reading 18 texts (the same amount of words per text) from Spain and Mexico that represent five diachronic cuts. Through reading full-texts, appearances of adjective-adverb interfaces are located without any lexical or formal pre-selection. Lemmatization and annotation of the discovered examples take into account a wider spectrum of adverb types, including prepositional phrases (*a la clara* 'clearly' *de seguro* 'surely', etc.), as well as a wider spectrum of adverbial functions: additionally to manner adverbs and the above mentioned categories, the corpus includes adverbial modificators (that is, specifiers and quantifiers) of nouns, adverbs and adjectives (for example: *muchos hombres harto sabios* 'many very wise men', *justo encima de las mesillas* 'just/exactly on the tables'), as well as sentence adverbs with discourse-functions (*cierto* 'truly'). Therefore, the categories "attribution target" (verb, verb and subject, noun, adjective, sentence. . .) and "semantics" (manner, time, place, discourse. . .) were added to the annotation model in order to categorize more information.

A second corpus on the diachrony of Spanish was compiled from records discovered during 2014 in the *Corpus del Diccionario Histórico*. This corpus (= Sp_A_CDH) contains examples from the 13th to the 21st century that were collected using a lemmata-list of the most frequent adjectives in Spanish. As a result, the database contains combinations of "*verb + adjective-adverb*". Thus, the data is limited to manner adverbs (e.g. *ver claro* 'to see clearly', *hablar alto* 'to speak loud', *correr rápido* 'to run fast', etc.). This corpus is also tagged according to morphology (e.g. inflection, as in *las piedras valen baratas* 'the stones cost cheap'), attribution target and semantics (e.g. manner, time, place, specification, ...).

2.2. Newly tagged data

During the project, new corpora were added to the database. Two corpora were collected and tagged by the project team:

A corpus based on Latin American Spanish examples from the 16th to the 19th century, discovered in the *Corpus Diacrónico y Diatópico del Español de América* (=CORDIAM), was elaborated during 2018 (=Sp_AP_CORDIAM). Examples of adjective-adverbs and prepositional phrases were collected by searching the lemma-list of adjective-adverbs discovered in the two former Spanish corpora. As a result, new combinations of "preposition + adjective" were discovered and tagged (for example, *por de pronto* 'provisionally, temporarily'). The annotation takes into account the wider spectrum of adverbial functions and meanings, as mentioned above.

The corpus of adjective-adverbs, *mente*-adverbs and prepositional phrases in 20th century Brazilian Portuguese (= Pt_APM_DeG) contains spoken and written language examples found in the corpus *Discurso & Gramática* (directed by de Oliveira & Votre at Universidade Federal do Rio de Janeiro in 1994). During the reading of the whole *Discurso & Gramática*-corpus (which is available as a full text on the website), the adverbs were annotated by the project team with the above mentioned wider categories. This enables systematic quantitative and qualitative analysis of triplets of different adverb-types (adjective-adverbs / *mente*-adverbs / prepositional phrases), such as *alto / altamente / por alto* ('highly') or *certo / certamente / ao certo* ('truly').

Two further corpora have been elaborated in cooperation with project partners by tagging their data using the annotation tool of the project:

The corpus *Adjective-adverbs in southern varieties of Italy* (=It_A_aaif) contains adjective-adverb examples collected and analysed by A. Ledgeway (University of Cambridge) for several research publications (for example, Ledgeway 2009). The collection is based on the reading of dialect literature from Naples (14th - 21st century), Sicilia, Calabria and Salento (20th - 21st century). These previously collected examples were annotated by the project team using the annotation tool. The corpus contains all types of adverbs (manner adverbs, specifiers, discourse markers, sentence adverbs, ...). Lemmatization is based on present-day standard Italian adjectives and verbs in order to unify different orthographic solutions of different regional varieties. The tagged category "reduplication" is especially interesting, since this phenomenon is widespread in southern Romance varieties of Italy, for example: *me parla doce doce dint"e rrecchie* 'she/he speaks sweet-sweet into my ears'.

The *Corpus of Old Romanian adjective-adverbs, derived adverbs and prepositional phrases* (= Ro_ADP_aaif) was collected and annotated in 2019 by A. Chircu (University of Cluj-Napoca). This corpus is based on the reading of ten texts from the 16th to 19th century (same extension for each extract). The full-text of the selected extracts are available online. The database includes adjective-adverbs (*drept* 'right', *foarte* 'strong') and prepositional phrases (*în scurt* or *pe scurt* 'shortly'). Since Romanian does not use *mente*-adverbs, but other derivational suffixes (namely *-ește, -iș,* and *-ul*, as in *cu dreptul* 'rightly, straightly'), the morphological categorization of the annotation tool has been extended. As the tagging is based on reading, there is no lexical or formal pre-selection and all types of adverbs are included.

3. Tokens, types and lemmata

Each adverb in the database was tagged with the lemma of the underlying adjective, e.g. all examples of the variants *alto / altamente / por alto* ('highly') are lemmatized as *alto*. The different formal structures of these types of adverbs (adjective-adverb, *mente*-adverb and prepositional phrase) are tagged separately under the category "morphosyntactic structure". Therefore, the lemmatization of the examples groups all morphologic variants to one basic adjective-lemma. Besides, the lemmatization unifies orthographic variation, especially regarding historical data.

Table 1 gives an overview of the annotated corpora. "Word tokens" refers to the total size of each corpus (some of them display the whole full-text), "tagged examples" refers to the total of annotated adverb-examples, and "adjective-lemmata" refers to how many different underlying adjectival lemmata were lemmatized in each corpus.

4. Annotation of linguistic categories and Data Modelling

Every example contains at least one annotated adverb and, depending on the respective corpus, the corresponding verb, preposition and/or the subject of the phrase. The model used for the annotation has been elaborated and extended during the creation of the various corpora by the research group. It has therefore proven useful for its practical application as well as for extensibility. The annotation model was developed to offer a cross-linguistic categorization model of the various forms, functions and meanings of the heterogeneous word-class "adverb" .

By using an annotation tool, every adjective-adverb example has been manually tagged by field-experts with several pre-defined morphosyntactic and semantic categories in order to ensure a consistent categorization of the examples. These categories complay the project-specific model, which in its current version is implemented as RDFs.[2] This ontology reflects the interdisciplinary research process and its iterative development during the project and ensures interoperability between the several resources (Pollin et al. 2018).

The annotation tool enables the manual categorisation according to this annotation model. Annotations are encoded in the *function* attribute of each token in the XML/TEI, which is the output format of the annotation tool. TEI represents a stable representation format for annotations, next to a big community that seeks for further development. The TEI can be used without the need to add further specifications and the possibility to modify and choose from a highly modular tag set (Stührenberg 2012). The decoding and semantic enrichment takes place in the processing phase, when the TEI/XML is ingested into the GAMS repository infrastructure. An RDF datastream is stored next to the TEI datastream in a single digital object and enables the use of data in the sense of Linked Open Data (Nordhoff & Hellmann 2012).

GAMS[3] is an OAIS compliant digital asset management system used for the administration, publication and long-term preservation of digital resources and enables scholars to publish resources. The repository, based on Fedora Commons, pursues a largely XML-based content strategy. A Blazegraph triplestore extends the infrastructure for the storage and aggregated querying of structured data like the language resources in the context of this project.

[2] AAIF-Ontology, gams.uni-graz.at/o:aaif.ontology

[3] Humanities' Asset Management System, http://gams.uni-graz.at

Corpus	Word tokens (total)	Tagged examples of adverbs	Types of adjective-lemmata	Main focus / adverb types	Coverage
Fr_A_DHAA	∼610.000	13.558	286	verbs + adjective-adverbs	11th - 20th c., France
Fr_A_Web	∼138.000	5.091	389	verbs + adjective-adverbs	21st c., France
Sp_AP_SH3	∼90.000	1.252	175	adjective-adverbs and prepositional phrases	13th–21st c., Spain & Mexico
Sp_A_CDH	∼81.000	2.424	161	verbs + adjective-adverbs	13th–21st c., Spain & Latin America
Sp_AP_CORDIAM	∼62.000	1.241	86	adjective-adverbs and prepositional phrases	16th–19th c., Latin America
Pt_APM_DeG	∼454.000 (full text displayed)	6.259	227	adjective-adverbs, *mente*-adverbs and prepositional phrases	20th c., Brazil
It_A_aaif	∼7.150	844	101	adjective-adverbs	14th–21stc., Naples / Sicilia, Salento, Calabria, 20th–21st c.
Ro_ADP_aaif	∼243.000 (full text displayed)	1.135	150	Adjective-adverbs, derived adverbs (*-ește, -iș, -ul*) and prepositional phrases	16th–19th c., Romania

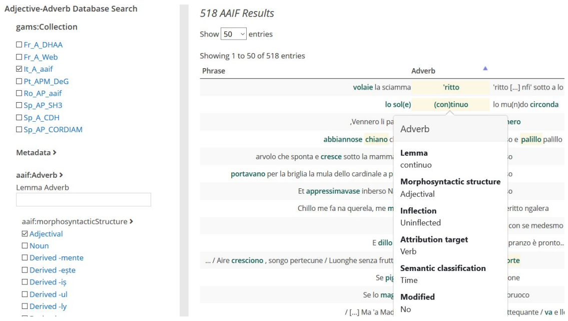
Table 1: Overview of the AAIF-Subcorpora

5. Database search

All corpora combined by the AAIF project are accessible over a dedicated web interface, which offers full-text access to the whole database, e.g. for entirely reading each one of the 8 subcorpora, including detailed bibliographic information about the sources of the texts. Additionally, the data can be queried via a web based search interface that offers fine-grained queries which can be a combination of the annotated categories/features and a lemma-search. As mentioned above, the lemmatization was done for the underlying adjective-lemma. Therefore, the lemma-search of a specific adjectival lemma displays all types of adverbs based on that adjective, including inflected adverbs, *mente*-Adverbs and prepositional phrases. This makes it possible to analyze type-token-frequencies.

The search mask also offers filtering of bibliographic metadata categories (author, year, region), as well as filtering for spoken or written data (only relevant for the Portuguese data). It is possible to search through only one corpus but also across a selection of several corpora as well as across a selection of languages.

Queries entered via the search interface are directly translated into a SPARQL query and forwarded to the triple store. The triples match the RDF, which is extracted from the TEI/XML (@function) and resolved as highly structured data according to the annotation model. Figure 1 shows the result of such a search query.[4] In this example, all entries from the corpus "It_A_aaif" are queried for which adverbs have an adjectival morphosyntactic structure and are not inflected. The result is displayed in a dynamic table and gives insight into the overall annotation of each adverb, as well as additional metadata about the evidence.



Adjective-Adverb Database Search

gams:Collection

☐ Fr_A_DHAA
☐ Fr_A_Web
☑ It_A_aaif
☐ Pt_APM_DeG
☐ Ro_AP_aaif
☐ Sp_AP_SH3
☐ Sp_A_CDH
☐ Sp_AP_CORDIAM

Metadata ›

aaif:Adverb ›
Lemma Adverb

aaif:morphosyntacticStructure ›
☑ Adjectival
☐ Noun
☐ Derived -mente
☐ Derived -ește
☐ Derived -iș
☐ Derived -ul
☐ Derived -ly

518 AAIF Results

Show 50 entries

Showing 1 to 50 of 518 entries

Phrase	Adverb	
volaie la sciamma	'ritto	'ritto [...] nfi' sotto a lo
lo sol(e)	(con)tinuo	lo mu(n)do circonda
,Vennero li pa		nero
abbiannose chiano c		o e palillo palillo
arvolo che sponta e cresce sotto la mamm		o
portavano per la briglia la mula dello cardinale a p		o
Et appressimavase inberso N		o
Chillo me fa na querela, me m		ritto ngalera
		con se medesmo
E dillo		pranzo è pronto..
... / Aire cresciono , songo pertecune / Luonghe senza frutt		rte
Se pi		one
Se lo mag		ruoco
/ [...] Ma 'a Mad		ttequante / va e ll

Adverb

Lemma
continuo

Morphosyntactic structure
Adjectival

Inflection
Uninflected

Attribution target
Verb

Semantic classification
Time

Modified
No

Figure 1: Search mask and display of results

Both the whole subcorpora and the results of queries in the search mask can be downloaded for further use in several formats: TEI/XML, RDF and Excel. In the Excel-Export of the whole subcorpora, the annotated categories are structured in columns. Furthermore, the project wants to ensure the expandability of the system: by providing Open Access to the annotation tool and by offering comprehensive descriptions and metadata, integrating new data via the GAMS-repository will still be possible once the project is finished.

[4] AAIF Search, gams.uni-graz.at/query:aaif.db

6. Conclusion and Outlook

In this paper, we discussed the work on an expandable database for adjective-adverb interfaces in Romance languages. The examples are manually selected and annotated with varying levels of depth. For every example at least one relevant adverb exists, which is annotated in regard of its morphosyntactic structure, attribution target and semantic classification. To be able to work with the data, we introduced a dedicated search interface that allows researchers to formulate structured queries to the individual subcorpora as well as jointly and across languages. This procedure is based on an annotation tool for manual tagging, whose XML/TEI-export is then converted into RDF. On the basis of this data, functionalities for searching and exporting language resources are offered on the project website. The database allows integrating new data from other research projects on the topic. We expect further datasets from an ongoing project in the research group that deals with adjectives as a part of adverbial locutions (prepositional phrases).[5] The annotation model has been created to serve as a standard model for classifying adverbs in several languages. Due to the similar usage of adjective-adverbs in other languages – not only Romance but also English – it is quite possible to open the database to those languages with little effort. In this respect, we aim to be open for data import from relevant projects and will provide the annotation tool as well as the annotation model under a free licence by the end of the project.

7. Bibliographical References

De Jong, F. M. G., Maegaard, B., De Smedt, K., Fišer, D., Van Uytvanck, D. (2018). CLARIN: towards FAIR and responsible data science using language resources. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), pp. 3259-3264.

Hummel, M. (2014). Los adjetivos adverbiales. In C. Company Company (ed.), *Sintaxis histórica de la lengua española.* México D.F.: Universidad Nacional Autónoma de México, pp. 615–733.

Hummel, M. (2017). Adjectives with adverbial functions in Romance. In M. Hummel & V. Salvador (eds.), *Adjective Adverb Interfaces in Romance*. Amsterdam: Philadelphia, pp. 13-46.

Hummel, M., Chircu, A., García Hernández, B., García Sánchez, J. J., Koch, S., Porcel Bueno, D. & Wissner, I. (2019). Prepositional Adverbials in the Diachrony of Romance: a State of the Art. *Zeitschrift für romanische Philologie* 135, fasc. 4, pp. 1-58.

Ledgeway, A. (2009). Grammatica diacronica del napoletano. Tübingen: Max Niemeyer Verlag.

Nordhoff, S., & Hellmann, S. (2012). Linked data in linguistics: Representing and connecting language data and language metadata. Berlin / Heidelberg: Springer.

Pollin, C., Schneider, G., Gerhalter, K. & Hummel, M. (2018). Semantic Annotation in the Project “Open Access Database ‘Adjective-Adverb Interfaces’ in Romance” . In Proceedings of the Workshop on Annotation in Digital Humanities. CEUR Workshop Proceedings, pp. 41-46.

Stigler J. & Steiner E. (2018). GAMS – An infrastructure for the long-term preservation and publication of research data from the Humanities. In: Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare. Mitteilungen, pp. 207-216.

Stührenberg, M. (2012). The TEI and current standards for structuring linguistic data. An overview. Journal of the Text Encoding Initiative 3.

8. Language Resource References

AAIF-Database = Schneider, G., Pollin, C., Gerhalter, K. & Hummel, M. (2019): Adjective-Adverb Interfaces in Romance. Open-Access Database. https://gams.uni-graz.at/context:aaif

Fr_A_DHAA = Hummel, M., Gerhalter, K., Schneider, G. & Pollin, C. (2019). Corpus French *Dictionnaire historique de l'adjectif-adverbe*. 2^{nd} version. AAIF-Database. https://gams.uni-graz.at/o:aaif.fradhaa

Fr_A_Web = Hummel, M., Telsnig, H., Korper, G., Gazdik, A., Höfferer, J. Gerhalter, K., Schneider, G. & Pollin, C. (2019). Corpus of French Adjective-Adverbs in Informal Texts from the Web. AAIF-Database. https://gams.uni-graz.at/o:aaif.fraweb

It_A_aaif = Ledgeway, A. & Gerhalter, K. (2019): Corpus of Adjective-Adverbs in Southern Varieties of Italy. In AAIF-Database. https://gams.uni-graz.at/o:aaif.itaaaif

Pt_APM_DeG = de Oliveira, A. R., Votre, S. J. & Gerhalter, K. (2019): Corpus of Brazilian Portuguese Adjective-Adverbs, *mente*-Adverbs and Prepositional Phrases in Corpus *Discurso* & *Gramática*. In AAIF-Database. https://gams.uni-graz.at/o:aaif.ptapmdeg

Ro_ADP_aaif = Chircu, A. (2019): Corpus of Old Romanian Adjective-Adverbs, Derived Adverbs and Prepositional Phrases. In AAIF-Database. https://gams.uni-graz.at/o:aaif.roapaaif

Sp_A_CDH = Gerhalter, K., Schneider, G., Pollin, C. & Hummel, M. (2019): Corpus of Spanish ”Verb + Adjective-Adverbs” in CDH. 2^{nd} version. AAIF-Database. https://gams.uni-graz.at/o:aaif.spacdh

Sp_AP_CORDIAM = Striedner, P., Gerhalter, K., Schneider, G., Pollin, C. & Hummel, M. (2019): Corpus of Spanish Adjective-Adverbs and Prepositional Phrases in CORDIAM. AAIF-Database. https://gams.uni-graz.at/o:aaif.spapcordiam

Sp_AP_SH3 = Gerhalter, K., Hummel, M., Schneider, G. & Pollin, C. (2019). Corpus of Spanish Adjective-Adverbs in Diachrony (for *Sintaxis Histórica III*). AAIF-Database. https://gams.uni-graz.at/o:aaif.spapsh3

[5] https://adjective-adverb.uni-graz.at/en/projects/the-third-way-2018-2021/