

# Using annotation and specialised electronic corpora to facilitate the reading of journal articles in the field of Law

Pascual Pérez-Paredes<sup>1</sup> (pascualf@um.es) María Sánchez-Tornel<sup>1</sup> (mstornel@um.es) José M. Alcaraz Calero<sup>2</sup> (jose.alcaraz-calero@hp.com)

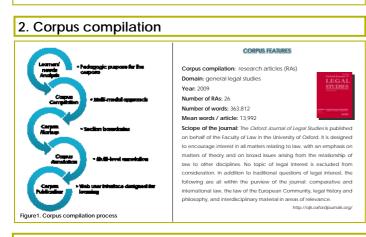
# 1:Universidad de Murcia 2: Cloud and Security Lab. Hewlett-Packard Research Laboratories

### Abstract

tering specialised languages is increasingly becoming a pressing need for protections in the academic is an effort to belance their training in specialised vocabulary acquisition and rea that many practice. The experience we report involves students seeking a PAD in Low of UA. Within their andiscitanting of the complexities of the students students students and reas ting. Hase learners need to go beyond maindmans English learning for general communication, their takes the development of learner shells of students students students and reas tang, index learners need to go beyond maindmans. English learning for general communication, their takes the development of learner shells of students. The copus candidad of 25 MAs p tang, contempleting reading and understanding research for their academic, scholar purposes. In this Ocdent Journal of Legal Studies and the total amound of words was 343,812. This poster to lear, scening to grips with the released students at academic (EAA) is of generic minimum for the students of the two of the two of the two of the students and the state students and the state students and the state students and the state students and the students and the total amound of words was 343,812. This poster to tell, scening to grips with the released students at reasons (EAA) is of generic minimum for the students of the two other expensions at the students attractions of the students and the total amound of words was 343,812. This poster to tell scenes (inguistic research stills.

# 1. Legal English for research purposes

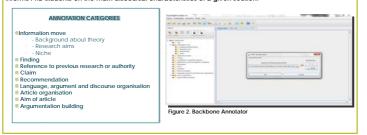
A comprehensive analysis of any professional genre must consider and integrate text-internal as well as text-external aspects of language use (Bhatia, 2004). Beyond the well-known difficulties of legal language (Bhatia, 1982) and its main features (Alcaraz Varó, 2002), the rhetorical structure of a RA is a reflection of the professional practices of scholars in the field of law. Linguists consider that a genre is an instance of a sociolinguistic activity through which members of certain discourse community achieve their communicative purposes (Swales, 1990). A genre is defined then by its shared communicative purpose and manifested by its particular structural and linguistic features. The Research Article (RA) is the main channel of scientific or scholarly communication, which turns it into the target genre for PhD candidates in law.



# 3. Corpus annotation

The corpus was marked and annotated using Backbone Annotator (3.8). BACKBONE Annotator is a tool that assists annotators in the process of manual annotation of corpora (Pérez-Paredes & Alcaraz, 2009). BACKBONE Annotator allows for the management of multiple corpus files and corpora. Annotators can create import or select different corpora and perform the same or different annotation schemes on each of them. Our annotation relies on three key aspects: (1) Annotation is unit-bound, not text bound. Our RAs are divided into paragraphs and then rhetoric moves

Annotation is unit-bound, not text bound. Our RAs are divided into paragraphs and then rhetoric moves are identified and annotated.
Annotation is based on the Text Encoding Initiative (TEI)-compliant XML. This gives our XML files the processing and structuring power of XML plus the uniformity and standardisation of TEI.
Our annotation approach has a multipurpose nature, in this case RAs rhetorical moves. The annotation informs PhD students on the main discoursal characteristics of a given section.



## Corpus availability

Once annotated, the resulting corpus was published on a server and learners (n=15) were provided with scaffolded guided-activities that sought to promote a problem-solving, data-driven learning approach. Learners were presented with activities that required using the annotated corpus and navigating the rhetorical moves therein.

Cadegory Ville	Assessed	Baartit mode	All of the selected categories	
🗢 🎠 📋 information move	(a) + Cold A4	Lass for	at sectors	
- Background about theiry	2	"Bawroth its revealing	B statistics	
Finding Reference to previous research or authority	a.	Benedit Cathopory Fill	12	
		Train the		
ethin that shows be the second state and show and the			Construction of the second	IN THE OF
ction as net once, or net syster, the trees reach together any, as essentiated or any rare in affairt balances has utiliant. He highly publicitud approach to	opera al caso o menor accura o e repayante transito, comerci ingenera, ve o etcla I CCTari in referenzi tu caso accurativamente Casona, anti a vener caso tra leat of proportionality referenzi tu casor development D atoma		адалын анд арсанда адалынун Коло	to fillo
<b>ction</b> As an approximate space part of the property state, we assume that the case of affect fundaments has utilized, the fulfying subperturbed approach to Convention sight, whether the definition is a public of a private book to	efforte & Chiltrat is referre (Ho sender development C above, and it every can		algoritem and decourse organization	10 HLD
CEAR International Contraction Network International Contractions of the Part of the Mercel Neurol Materia, the Nythy Audio-Mercell Approach to Constrain Hyper Contraction Network International Contraction and architecture (EN Text Approx 4 specificity with which contract) (Mercels should be Textme Andresets, Text Approximation, In 1999 Text Approx 2004) and Externation Contractions, Text Approximation, In 1999 Text Approximation Science Science Contra- tioned Sciences, Text Approximation, In 1999 Text Approximation, International Approximation, International Approximation, International Approximation, International Science Contractions, Development, Contraction, International, International Approximation, International Approximation	which ECTION is reducing to under development C alone, and it were) can be let of proportionally informatic under development D alone and shouldned that been the autject of much attention in scantyware both development and of the evelop property directs that, social	e extension the enterterance with an annual sector to the sector to the sector to the sector of the	algement and decorded organization forces for regulating	10 1110
If an is addit balances has balance. If a highly plaquement approach to Consolition right called within the definition is a public of a public body. The degree of parenticity with which consolition and the term degrees. The average of parenticity within the called balance and degrees of the average is in 1995 the called balance contrastence place and degrees of the average is in 1995 the called balance and degrees of the average is integret and balance and degrees of the average is in 1995 the called balance and degrees of the average is integret and the average of the place and their integreement is a save degree and the off and a new degree average and their integreement is a save and there a the out of a new degree and their integreement is a save and there a the out of a new degree and their integreement is a save and the save and the output of the save and their integreement is a save and there a theory of the save and their integreement is a save and there in the save and the save and their integreement is a save and the save and the save and the save and the save and the save and the save and the save and the save and the save and th	which is Carl to index and to an end of the second and the second	In outputs the markening with the second sec	algomens and docord al ingencement Set mynighting feer	10 1110
chain any in Advances to be been more separations, as estimates in the pay in Advances to Balances the Statistical Statistical Statistical Constraining of Advances to Balances to Balances and Statistical end in Yanakes (Ed.) The days of a second statistical Constraint Advances should be formed detects of years and a second statistical Statistical Statistical Advances and the second statistical Statistical Statistical Statistical Advances and the second statistical Statistical Statistical Statistical Advances and Statistical Statistical Statistical Statistical Statistical Statistical Advances and statistical Statistic	disk to Clinical investments to under intervention of unions, and its revel tase for that of programmental index to entry thereing memory and and structures of the union of the entry directory of unions, and con- sent structures of the union of the entry directory of the union and the union of the union of the union of the union of the union of the union of the union of the union of the union of the union of the union of the union of the union of the union of the union of the union of the union of the union	a publicità the attentivence with tra publicità di anti- tra care interne suddes and conserve de la conserve de la conserve suddes attenti de la conserve de la conserve suddes attenti de la conserve de la conserve tenti de la conserve de la conse	argament and opposited organization bet représent	

### 5. Learners' evaluation & conclusion

Students (n=7) completed a questionnaire once they submitted their answers. Learners found the resource useful (Q1) and said that reading concordances (Q2) and having access to annotated resources facilitates their understanding of hetorical moves in RAs (Q3). The evaluation of the learners seems to be in line with line previous findings on using corpora for the teaching of specialised languages (Fan & Xunfeng, 2002; Yoon & Hirvela, 2004).

Q1. I found this resource very useful as an aid for reading and comprehension
Q2. Concordance lines helped me to better understand the use of lexis
Q3. Concordance lines helped me to better understand the different moves in research articles
Q4.This type of resource is very useful for the study of Legal English
Q5. If you had the chance of using this resource for free, would you use it again?

ding, as well as ih

shed in 20

w Lás w

### References

Alcaraz Varó, E.(2007). El inglés jurídico. Barcelona: Ariel

hatia, V. K., (1982): An investigation into the formal and functional characteristics of qualifications in Legislative writing and its application to English for academic legal purposes. (Doctoral dissertation). University of Aston, Birmingham, U.K. Bhatia, V.K. (2004). Worlds of written discourse: a genre-based view. London: Continuum International

Fan, M. & Xunfeng, X. (2002). An evaluation of an online bilingual corpus for the self-learning of legal English. System (30)1, 47-63 Pérez-Paredes, P. & Alcaraz-Calero, J.M. (2009). Developing annotation solutions for online data driven learning. ReCALL (21)1, 55-75. Swales, J. M. (1990). Genre analysis: English in academic and research settings. Cambridge: Cambridge University Press Yoon, H., & Hirvela, A. (2004). ESL student attitudes towards corpus use in L2 writing. Journal of Second Language Writing(13) 4, 257 283.