

CURRICULUM VITAE ABREVIADO (CVA)

IMPORTANT – The Curriculum Vitae cannot exceed 4 pages. Instructions to fill this document are available in the website.

Part A. PERSONAL INFORMATION

First name	Gloria		
Family name	Corpas Pastor		
Gender (*)	Female	Birth date (dd/mm/yyyy)	13/04/1965
Social Security, Passport, ID number	25078844-N		
e-mail	gcorpas@uma.es	URL Web https://www.lexytrad.es/es/miembros/gloria-corporas-pastor/	
Open Researcher and Contributor ID (ORCID) (*)	0000-0001-6688-1531		

(*) Mandatory

A.1. Current position

Position	Professor in Translation and Interpreting		
Initial date	2008		
Institution	University of Malaga		
Department/Center	Department of Translation and Interpreting	Faculty of Arts	
Country	Spain	Teleph. number	617 595551
Key words	Translation and Interpreting Technologies, Phraseology, Corpus Linguistics, Lexicography		

A.2. Education

PhD, Licensed, Graduate	University/Country	Year
B.A. and M.A. in English Philology	Universidad de Málaga	1988
Ph.D. in English Philology	Universidad Complutense de Madrid	1994
M.A./Expert degree in Internet and its Applications	Universidad Nacional de Educación a Distancia-UNED	2002

Part B. CV SUMMARY (max. 5000 characters, including spaces)

General indicators of the research quality

Prof Corpas has a total of **Six-Year Research and Transfer Periods** (five research periods and one transfer period) granted by the Spanish Ministry of Education, and six **Research, Teaching and Administration Periods** granted by the Andalusian Regional Ministry of Education, in addition to **six teaching periods** granted by the University of Malaga. At present, she has a total of 8739 citations in Google Scholar (h-index: 38; i-10 index: 113). In the last five years, her publications have been cited 3201 times (h-index: 28; i-10 index: 62).

Prof Corpas Pastor has supervised a total of 28 PhD theses: 21 (75%) cover language technologies (machine translation, machine interpretation, CAT tools, CAI tools, corpus linguistics, natural language processing, etc.), of which 7 (33.33%) focus on tourism. She is currently supervising 8 PhD thesis on various aspects of multilingual language technologies (2 of them near to completion).

CV summary

BA in German Philology (English) from the University of Malaga. PhD in English Philology from the Universidad Complutense de Madrid (1994). Visiting Professor in Translation Technology at the Research Institute in Information and Language Processing (RIILP) of the University of Wolverhampton, UK (2007-2023), honorary Adjunct Professor at Xi'an Jiaotong-Liverpool University, China (since 2020), and Professor in Translation and Interpreting at the University of Malaga, Spain (since 2008). Published and cited extensively, member of several international and national editorial and scientific committees. Spanish delegate for AEN/CTN 174 and CEN/BTTF 138, Spanish expert for TC37/SC5, where she has been convenor of ISO TC37/SC5AHG2 "Speech to Text" (2019-2023). Extensive experience in institutional evaluation programmes, including verification, monitoring and accreditation (ANECA, AQU, ACCUEE, DEVA, AQU, ACCUA, etc.). President of the Evaluation and Verification Commission (Arts and Humanities) at Fundación para el Conocimiento Madrid, and President of the Evaluation Committee (the Arts and Humanities) at ACCUEE. She also collaborates regularly as evaluator for the Spanish Ministry in various R&D and youth employment programmes. She has recently been awarded the "2017 Farola Award" in the field of New Technologies, given by the Andalusian Women's Institute, Junta de Andalucía; the Certificate of Commendation-Saint Francis Prize in Techno-Humanities 2022, given by the Caritas Institute of Higher Education, Hong Kong; the "Doctora de Alcalá Award" for her excellent track-record in Arts and the Humanities (2022), given by the University of Alcalá; "Flag of Andalucía Award" on Research, Science and Healthcare (2023), granted by the Andalusian Government, the "Carlomagno Research Award" in Arts and Humanities (2023) and First Prize in the Spin-Off competition (XXVII Edition, 2023) and II Edition of the University of Malaga Award for Knowledge Transfer (2024). National Research Prize (Ramón Menéndez Pidal modality), awarded by the Spanish Ministry of Science, Innovation and Universities (2024). President of AIETI (Iberian Association of Translation and Interpreting Studies, 2015-2017), Vice-President of AMIT-A (Association of Women in Science and Technology of Andalusia, 2014-2017), and Director of the Department of Translation and Interpreting of the University of Malaga (2016-2021). She is currently director of the Research Institute in Multilingual Language Technologies (IUITLM).

Interdisciplinary profile

Gloria Corpas Pastor has an extensive scientific background in language technologies applied to translation and interpreting, and a double research profile in both Translation and Interpreting and NLP and AI.

The research projects that she is currently leading or has led in the past are evidence of the interdisciplinary nature of her research. The ultimate goal of the project "Multilingual dialogue systems using neural networks for apps in the healthcare domain: the triage (Spanish-English/Arabic)" [4], for example, was the development of a multilingual system to automate triage, designing and implementing a multilingual system to enable effective communication between the healthcare professional and the patient. The methodology used combined various Natural Language Processing techniques, such as neural-based machine translation, machine learning, speech recognition (speech-text) and synthesis (text-to-speech), in addition to translation and interpretation, corpus linguistics, digital resources and multilingual linguistic technologies.

Another project, the VIP system (VIP II) [1], also reflects this multidisciplinary by combining corpus linguistics with natural language processing to revolutionise the interpreting industry by creating a set of tools for interpreters.

Part C. RELEVANT MERITS (sorted by typology)

C.1. Publications (see instructions)

1. Lima Florido, F. J., & Corpas Pastor, G. 2025. Advanced Identification of Prosodic Boundaries, Speakers, and Accents Through Multi-Task Audio Pre-Processing and Speech Language Models. *Computers*, 14(3), 102. <https://doi.org/10.3390/computers14030102>.

2. Corpas Pastor, G. 2024. "Ready for the Future? Towards the "Tech" Revolution in Interpreting". *Routledge Encyclopedia of Technology and the Humanities*, C. Sin-Wai (ed.). London: Routledge. 183-199.
3. Corpas Pastor, G. & F. J. Veredas Navarro (eds.). 2024. *Tecnologías Lingüísticas Multilingües. Desarrollos actuales y transición digital*. Granada: Comares. ISBN: 978-84-1369-807-6.
4. Lima Florido, F. J. & Corpas Pastor, G. 2024. "Posibilidades (y limitaciones) de los asistentes conversacionales como recursos tecnológicos para la interpretación: ChatGPT y Gemini frente a VIP". En *Tecnologías Lingüísticas Multilingües. Desarrollos actuales y transición digital* (eds. Corpas Pastor, G. & F. J. Veredas Navarro). Granada: Comares. 07–40. ISBN: 978-84-1369-807-6.
5. Lancheros, B. S.; Corpas Pastor, G. & R. Mitkov. 2024. "Data augmentation and transfer learning for cross-lingual named entity recognition in the biomedical domain". *Language Resources and Evaluation*, 2(58). ISSN: 1578-7559. DOI: <https://doi.org/10.1007/s10579-024-09738-8>.
6. Monti, J.; Corpas Pastor, G, Mitkov R. & C.M Hidalgo-Ternero (eds.). 2024. *Recent Advances in Multiword Units in Machine Translation and Translation Technology*. John Benjamins Publishing Company. ISBN: 9027246386. E-ISBN: 9789027246387. DOI: <https://doi.org/10.1075/cilt.366>.
7. Corpas Pastor, G. & B. Defrancq (eds.). 2023. *Interpreting Technologies. "IVITRA Research in Linguistics and Literature"*. Amsterdam: John Benjamins.
8. Bowker, L. & G. Corpas Pastor. 2022. "Translation Technology". In *The Oxford Handbook of Computational Linguistics*. Second edition (ed. R. Mitkov). Oxford: Oxford University Press. 871-905. ISBN: 0199573697, 978-0199573691.

C.2. Research projects

1. JA. A1.3-06, **DÍGAME: Diseño y proGrAmación de un prototipo multilingüe para la comunicación en enferMERía de urgencias y emergencias [Design and programming of a multilingual prototype for emergency nursing communication]**. European Regional Development Fund (ERDF), within the Framework of the II own Research, Transfer and Dissemination Plan of the University of Malaga. PI: G. Corpas. € 91 200.
2. ProyExcel_00540, **RECOVER: Multilingual machine interpretation for COVID-19 cases in emergency departments**. Projects of Excellence, Andalusian Regional Government (Junta de Andalucía), PAIDI 2020. PI: G. Corpas. 2022-2025. € 79 994.71.
3. PID2020-112818GB-I00, **Multi-lingual and Multi-domain Adaptation for the Optimisation of the VIP system (VIP II)**. Spanish Ministry of Science and Innovation. PI: G. Corpas. 2021-2025. € 60 500.
4. PDC2021-121220-I00, **Voice-text integrated system for InterPreters: Proof of Concept**. Spanish Ministry of Science and Innovation. PI: G. Corpas. 2021-2023. € 92 000.
5. CA19102, **LITHME: Language in the Human-Machine Era**. Cost action Horizonte 2020. PI: Dave Sayers (University of Jyväskylä, Finland).
6. UMA18-FEDERJA-067, **Multilingual dialogue systems using neural networks for apps in the healthcare domain: the triage (Spanish-English/Arabic)**. European Regional Development Fund (ERDF), under the Operational Programme "Andalusia". PI: G. Corpas. 2019-2022. € 34 401,53.
7. 599287-EPP-1-2018-1-UK-EPPKA1-JMD-MOB, **European Masters' in Technology for Translation and Interpreting**. European Commission. Erasmus+. Higher education - Erasmus Mundus Joint Master Degree. Consortium: University of Wolverhampton (UK, Coordinator), University of Malaga (Spain, G. Corpas is the PI of UMA), New Bulgarian University (Bulgaria) and Ghent University (Belgium). 2018-2024. € 3 052 000.
8. EQC2018-004572-P, **Application of Advanced NLP Techniques to the Field of Translation and Interpreting Technologies**. Spanish Ministry of Economy and Competitiveness. National Sub-Programme for Scientific and Technological Infrastructures and Equipment. PI: G. Corpas. 2018-2020. € 124 744.78.

9. UMA-CEIATECH-04, **MI4ALL - Machine Interpretation for All Through a Deep Learning API**. Innovation, Science and Enterprise Council of the Andalusian Regional Government through the Campus of International Excellence-Andalucía Tech. Singular project for knowledge transfer actions. PI: G. Corpas. 2019-2021. € 31 169.70

C.4. Contracts, technological or transfer merits, Include patents and other industrial or intellectual property activities (contracts, licenses, agreements, etc.) in which you have collaborated. Indicate: a) the order of signature of authors; b) reference; c) title; d) priority countries; e) date; f) Entity and companies that exploit the patent or similar information, if any.

C.4.1. Contracts

1. **Contract:** Participation as member of the Echegaray Committee, associated to the "Echegaray Programme". *Fundación para el Conocimiento Madri+d* (Contract no. 8.06/5.73.5529 through the OTRI Department, University of Malaga). 06/02/2020-06/06/2022. €18 000.
2. **Contract:** Technical and academic consultancy for the company Higher Ed Partners Limited (Contract no. 8.06/6.33.5879 through the OTRI Department, University of Malaga). 19/05/2021- ongoing. € 25 000.
3. **Contract:**, Scientific consultant for the Development of Scientific, Technical or Artistic Activities between the University of Malaga and the International University of La Rioja (Contract under Art. 60 of the Organic Law of the University System (LOSU). 2021-ongoing.

C.4.2. Patents and Intellectual Property Registries

1. **Patent:** G. Corpas, M. Seghiri, R. Maggi. *Nuevo método para determinar la representatividad de un corpus* [A New Method for Determining Corpus Representativeness]. Spanish Patent Office. Ref. no. P200603157. 03/02/2010. University of Malaga.
2. **Intellectual Property Registry:** G. Corpas Pastor, F. J. Lima Florido. Intellectual Property Registry of the tool "Herramienta de asistencia multilingüe para Pre-triage de COVID-19". SAFE CREATIVE. Registration code: 2203170739055. 2022. University of Malaga.
3. **Intellectual Property Registry:** G. Corpas, J. A. Fernández, F. J. Lima. Intellectual Property Registry of the CORPUS, GLOSSARY, COMPLEMENTARY and TRAINING packages of the VIP system. SAFE CREATIVE. 2021. University of Malaga.
4. **Intellectual Property Registry:** C. Hidalgo, G. Corpas, J. A. Fernández. Intellectual Property Registry of *gApp*. SAFE CREATIVE. 2020. University of Malaga
5. **Intellectual Property Registry:** G. Corpas, F. J. Lima, J. A. Fernández. Intellectual Property Registry of the *Sistemas de diálogo multilingües de base neuronal para el desarrollo de apps en el ámbito sanitario - API*. SAFE CREATIVE. Registration code: 2312306515712. 2023. University of Malaga.

C.4.3 Resources

Resources resulting from these projects and other research coordinated by G. Corpas::

1. **Multilingual intelligent dictionaries**, such as [INTELITERM](#) and [TERMITUR](#).
2. A **tool for interpreters**: the [VIP](#) system, which offers assistance to interpreters before, during and after the interpretation. It integrates terminology tools, corpus compilation and processing, automatic glossary creation, automatic speech recognition and quality assessment applications.
3. A [multilingual system](#) to automate triage, as well as a [COVID pre-triage system](#).
4. An application, [gApp](#), which improves neural machine translation in its application to **phraseology**.
5. **Terminology applications**, such as [TRANDIX](#), [OntoDiccionario](#) and [SHIPSAFE](#).