

Jerid C. Francom

Winston-Salem, NC

✉ francojc@wfu.edu  [francojc](#)  0000-0001-5972-6330 | Updated: Feb. 13, 2025

Experience

Associate Chair <i>Department of Spanish and Italian, Wake Forest University</i>	2019-2022
Program Director <i>Interdisciplinary Linguistics Minor, Wake Forest University</i>	2016-2019
Associate Professor <i>Department of Romance Languages, Wake Forest University</i>	2015-present
Research Consultant <i>Center for the Advanced Study of Language, University of Maryland</i>	2012-2013
Assistant Professor <i>Department of Romance Languages, Wake Forest University</i>	2009-2015

Education

PhD Linguistics & Cognitive Science <i>University of Arizona</i> Co-chairs: Janet Nicol & Simin Karimi Dissertation: <i>Experimental Syntax: exploring the effect of repeated exposure to anomalous syntactic structure – evidence from rating and reading tasks.</i> [link]	2004-2009
MA Hispanic Linguistics <i>University of Arizona</i> Chair: Antxon Olarrea Thesis: <i>Adjective small clauses and free-word order in Spanish</i> [link]	2001-2004
Certificate in English Language Teaching to Adults (CELTA) <i>University of Cambridge</i>	1998
BA History <i>University of Arizona</i> Outstanding research paper: <i>Crime and castigation —prohibition in Tucson</i>	1996-1998

Scholarship

268 total citations
h-index 8

BOOKS

Francom, J. (2025). *An Introduction to Quantitative Text Analysis for Linguistics: Reproducible Research Using R*. Routledge. Retrieved from <https://www.routledge.com/An-Introduction-to-Quantitative-Text-Analysis-for-Linguistics-Reproducible-Research-Using-R/Francom/p/book/9781032494265>

BOOK CHAPTERS

Francom, J. (2022). Corpus studies of syntax. In G. Goodall (Ed.), *The Cambridge Handbook of Experimental Syntax* (pp. 687–713). Cambridge University Press.

Hulden, M., & Francom, J. (2016). Spanish Diacritic Error Detection and Restoration—A Survey. In Z. Vetulani, H. Uszkoreit, & M. Kubis (Eds.), *Lecture Notes in Computer Science* (Vol. 9561, pp. 290–303). Springer. <https://doi.org/10.1007/978-3-642-20095-3>

Francom, J. (2011). Wh-movement: Interrogatives, exclamatives and relatives. In J. I. Hualde, A. Olarrea, & E. O'Rourke (Eds.), *The Handbook of Hispanic Linguistics* (pp. 533–556). Wiley-Blackwell.

REFEREED ARTICLES

Hulden, M., Silfverberg, M., & Francom, J. (2015). Kernel density estimation for text-based geolocation. In *Proceedings of the twenty-ninth meeting of the Association for the Advancement of Artificial Intelligence (AAAI)* (p. 6).

Francom, J., Hulden, M., & Ussishkin, A. (2014). ACTIV-ES: A comparable, cross-dialect corpus of “everyday” Spanish from Argentina, Mexico and Spain. In *The ninth international conference on Language Resources and Evaluation* (pp. 1733–1737).

Francom, J., & Hulden, M. (2013). Diacritic error detection and restoration via part-of-speech tags. In *Proceedings of the 6th Language and Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*. Poznan, Poland.

Hulden, M., & Francom, J. (2013). Weighted and unweighted transducers for tweet normalization. In *Tweet Normalization Workshop at SEPLN 2013*. Madrid, Spain: Sun SITE Central Europe (CEUR) Workshop Proceedings. Retrieved from <http://ceur-ws.org/Vol-1086/>

Hulden, M., Silfverberg, M., & Francom, J. (2013). Finite state applications with Javascript. In *Proceedings of the 19th Nordic Conference of Computational Linguistics* (Vol. 2, pp. 441–446).

Hulden, M., & Francom, J. (2012). Boosting Statistical Tagger Accuracy with Simple Rule-Based Grammars. In *Proceedings of the Eighth International Language Resources and Evaluation (LREC'12)* (pp. 2114–2117).

Francom, J., LaCross, A., & Ussishkin, A. (2010). How specialized are specialized corpora? Behavioral evaluation of corpus representativeness for Maltese. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC '10)*. Valletta, Malta: European Language Resources Association (ELRA).

Francom, J., Ussishkin, A., & Lacross, A. (2010). How specialized are specialized corpora?, 7.

Francom, J., Ussishkin, A., & Woudstra, D. (2009). Creating a Web-based Lexical Corpus and Information-extraction Tools for the Semitic Language Maltese. In *Proceedings of the SALTMIL 2009 workshop on Information Retrieval and Information Extraction for Less Resourced Languages*. (pp. 9–16).

Francom, J. (2008). Is Lexical Access Mediated by the Syllabic and/or CV Structure of Words? Exploring Transposed-Letter Priming Effects. In *UW Working Papers in Linguistics*.

Francom, J., & Hulden, M. (2008). Parallel Multi-Theory Annotations of Syntactic Structure. In E.

- L. R. A. (ELRA) (Ed.), *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)* (pp. 2339–2343). Marrakech, Morocco.
- Francom, J. (2003a). La cláusula reducida y el libre orden de palabras en español. *Divergencias: Revista de Estudios Lingüísticos y Literarios*, 1(1).
- Francom, J. (2003b). Variación estilística y dialecto: Un estudio de acomodación de lenguaje en el suroeste de Arizona. In *Annual Proceedings of the Thirteenth Annual Hispanic Symposium*. Tucson: University of Arizona Press.

PRESENTATIONS

- Francom, J. (2014a). *A machine learning approach to high(er) resolution mapping of linguistic variation*.
- Francom, J. (2014b). *Data analytics and language jobs*.
- Francom, J. (2014c). *Exploring voting-based QA systems Wake Forest University*.
- Francom, J. (2014d). *Towards high(er) resolution language variation mapping using social media*.
- Francom, J., Hulden, M., & Ussishkin, A. (2014). *ACTIV-ES: A comparable, cross-dialect corpus of “everyday” Spanish from Argentina, Mexico, and Spain*. Poster presented at the Language Resources and Evaluation Conference (LREC 14), Istanbul, Turkey.
- Francom, J. (2013a). Building a comparable corpus of “everyday” Spanish in Argentina, Mexico, and Spain.
- Francom, J. (2013b). Using the ACTIV-ES corpus to enhance Spanish language learning.
- Francom, J. (2012a). Hands on with Mendeley. Winston-Salem, NC: Mendeley Workshop.
- Francom, J. (2012b). The ACTIV-ES corpus: Moving towards cross-dialectal, ecologically validated language samples, and a more inclusive language science. Arizona Linguistics Circle 6.
- Hulden, M., & Francom, J. (2012). *Boosting Statistical Tagger Accuracy with Simple Rule-Based Grammars*. Presented at the Language Resource and Evaluation Conference, Istanbul, Turkey.
- Francom, J. (2011a). Are bilinguals two monolinguals? Raleigh, NC: Spanish Linguistics in North Carolina.
- Francom, J. (2011b). Mendeley: The cloud-computing answer to discovering, organizing and citing academic research in the digital age. Winston-Salem, NC: Mendeley Workshop.
- Francom, J. (2011c). SpanMorph and SpanSyn: Open-source language resources for the morphological and syntactic analysis of Spanish. Winston-Salem, NC: Linguistics Circle Meeting.
- Francom, J. (2011d). SpanMorph: An open-source morphological analyzer based on finite-state technology. Winston-Salem, NC: Linguistics Circle Meeting.
- Francom, J. (2011e). What is digital humanities? Winston-Salem, NC: WFU Digital Humanities Mixer.
- Francom, J., & Ussishkin, A. (2011a). Converging methodologies: Assessing corpus representativeness through psycholinguistic measures. Atlanta, Georgia: American Association for Corpus Linguistics.
- Francom, J., & Ussishkin, A. (2011b). Predicting word frequency effects from corpora: An investigation of corpus size and sampling methods. Pittsburgh, PA: Linguistic Society of America.
- Francom, J. (2010). Developing a Spanish active-language corpus for applications in the humanities and beyond. Winston-Salem, NC: Linguistics Circle Meeting.
- Francom, J., Ussishkin, A., & Lacross, A. (2010). How specialized are specialized corpora?, 7.
- Francom, J., Ussishkin, A., Lacross, A., & Meeting, L. C. (2010). How specialized are specialized corpora? Behavioral evaluation of corpus representativeness for Maltese. Winston-Salem, NC: Linguistics Circle Meeting.
- Hulden, M., Francom, J., & Tubino-blanco, M. (2010). SpanMorph: A open-source morphological analyzer for Spanish. Hermosillo, Mexico: Universidad de Sonora.

- Francom, J. (2009a). Assessing Experimental Syntax: Data from rating and reading time measures. Winston-Salem, NC: Linguistics Circle Meeting.
- Francom, J. (2009b). Is Syntactic Facilitation Contingent on Licit Syntactic Structure? San Francisco, CA: Linguistic Society of America.
- Francom, J., Woudstra, D., & Ussishkin, A. (2009a). Creating a Web-based Lexical Corpus and Information-extraction Tools for the Semitic Language Maltese. San Sebastian, Spain: Sociedad Española para el Procesamiento de Lenguaje Natural.
- Francom, J., Woudstra, D., & Ussishkin, A. (2009b). Electronic Corpora for Two Semitic Languages. Edmonton, Canada: American Association for Corpus Linguistics.
- Francom, J. (2008a). Is Lexical Access Mediated by the Syllabic and/ or CV Structure of Words? Seattle, WA: University of Washington.
- Francom, J. (2008b). Is Syntactic Satiation Contingent on Judgment-based Exposure? Hermosillo, Mexico: Universidad de Sonora.
- Francom, J. (2008c). Lexical & Syntactic Distributional Properties of Relatives in a Large Corpus. Tucson, AZ: Arizona Linguistics Circle 3.
- Francom, J., & Hulden, M. (2008). Parallel multi-theory annotation of syntactic structure. Marrakech, Morocco: Language Evaluation; Resources Conference.
- Francom, J. (2006). Splitting the Split Topic Analysis : Refining the Left Periphery in Spanish. Modern Language Association.
- Francom, J. (2003a). Modern Human Language: Learned or innate? Tucson, AZ: Spanish & Portuguese Symposium.
- Francom, J. (2003b). The Role of Community in Language Maintenance and Shift in Barcelona, Spain. Tucson, AZ: TINKER.

Mentoring

UNDERGRADUATE

Honors thesis advisees

2016. Madeline Nelson. ‘Linguistic Revitalization: the case of Catalan’

GRADUATE

Thesis advisees

2021. Dustin Rosh. ‘An Investigation in the Sociolinguistic Landscape and Diversity of Spain’
2017. Ying Ma. ‘Explicitation via the use of connectives in English-Chinese translation: A corpus-based study’

Grants

EXTERNAL

2017. National Science Foundation (NSF). *Improving Scientific Writing In Undergraduate STEM Classrooms: A Training Program for Students and Teaching Assistants Aided By Information Extraction Technology* Co-Primary Investigator. (1712423)

- 2011. National Endowment for the Humanities (NEH). *ACTIV-ES: a novel Spanish language corpus for linguistic and cultural comparisons between communities of the Hispanic world* Primary Investigator. ([HD 51432-II](#))
- 2003. Tinker Foundation Incorporated (TFI). Field Research Grant. Research in Barcelona, Spain. Primary Investigator. *The Role of Community in Language Maintenance and Shift in Barcelona, Spain*
- 2002. U.S. Department of Education. Foreign Language and Area Studies (FLAS) Fellow. Universidade Federal de Ceará, Fortaleza, Brazil.

INTERNAL

- 2016. Reynolds Research Leave. Office of the Dean. *Data Science for Linguists: concepts through code in R* Wake Forest University.
- 2016. Mid-Career Faculty Development Program. Office of the Dean. Wake Forest University.
- 2015. Office of the Provost's Summer Grants for Exploration of Educational Technology. *Student-directed composition assessment using web-based text analysis in RStudio's Shiny* Wake Forest University.
- 2014. Office of the Provost's Summer Grants for Exploration of Educational Technology. *Novel perspectives in language research and teaching through 'big data' and high-performance computing* Wake Forest University.
- 2013. Archie Fund for the Arts and the Humanities Grant. *Evaluating the ACTIV-ES corpus: a Spanish language corpus for three linguistically, culturally, and geographically distinct communities of the Hispanic world* Wake Forest University.
- 2013. Office of the Provost's Summer Grants for Exploration of Educational Technology. *Exploring voting-based QA systems to build out-of-class learning communities and robust student-centered knowledge-bases* Wake Forest University.
- 2011. Office of the Provost's Summer Grants for Exploration of Educational Technology. *Creating Highly Interactive Classroom Experiences in a Distance-learning Paradigm* Wake Forest University.
- 2011. Archie Fund for the Arts and the Humanities Grant. *SpanMorph and SpanSyn: an international collaborative effort to create open-source language resources for the morphological and syntactic analysis of Spanish* Wake Forest University.
- 2010. Office of the Provost's Summer Grants for Exploration of Educational Technology. *Real-world Examples in Real Time* Wake Forest University.

Professional development

- 2014. Lisbon Machine Learning School (LxMLS). Instituto Superior Técnico, Instituto de Telecomunicações. Lisbon, Portugal. July 22-29, 2014
- 2013. Corpus Linguistics and Statistics for Linguists with R. Georgia State University, Atlanta, GA. July 31-August 14, 2013