Text and Data Mining Panel Discussion

ProQuest Dissertations & Theses

CGS 62nd Annual Meeting| December 8, 2022

ProQuest

Part of **Clarivate**

Global Scholarship by the Numbers

5M+

Works, including 2.9M in Full text

4.1K

Institutional contributors from around the world; dissertations and theses in **64 languages**

240K+

Works added in 2022

400+

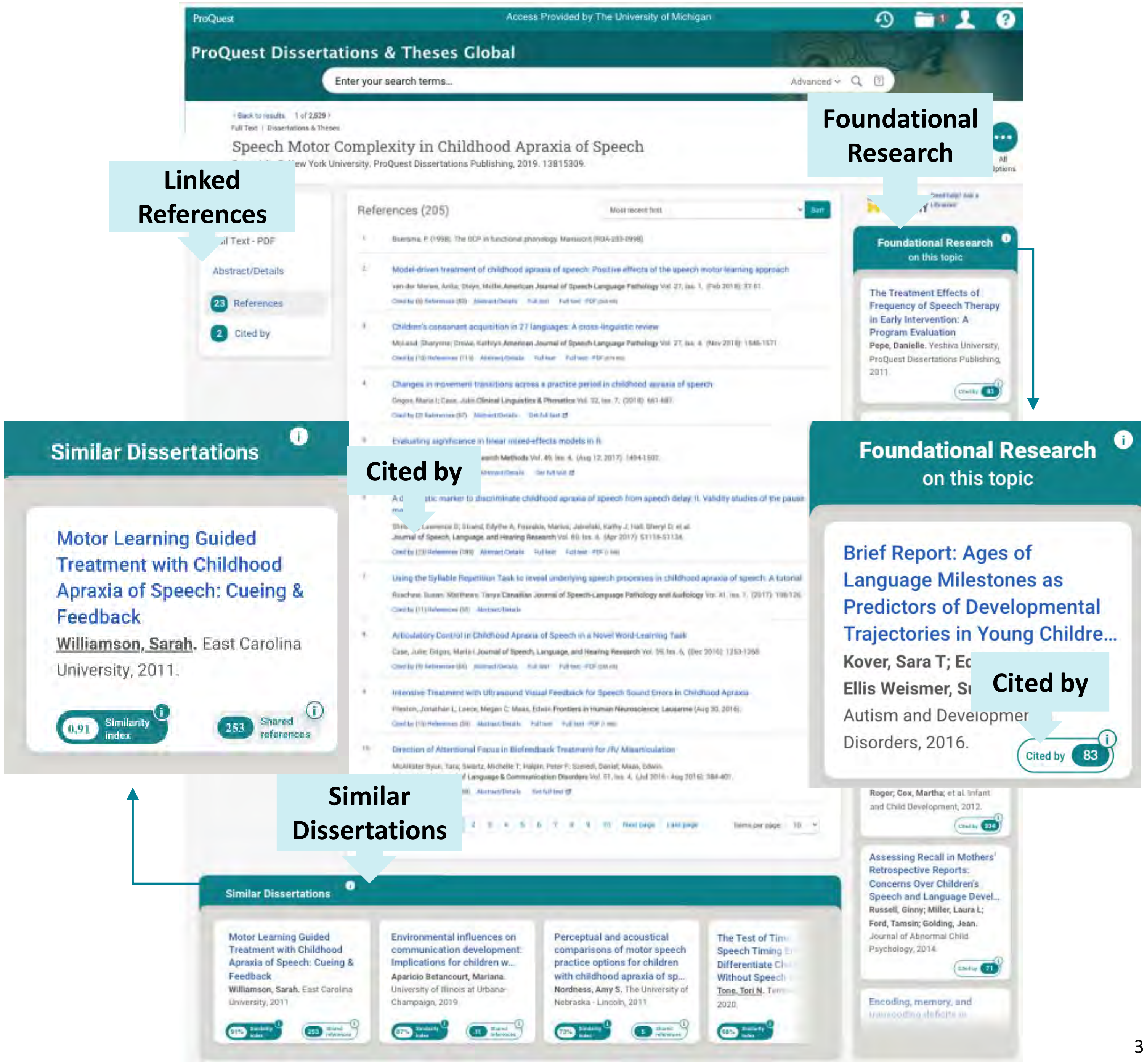
Years of coverage, from 1637 to 2022

ProQuest. ExLibris. Part of Clarivate

Citation Connections

Dynamic algorithms surface relevant, credible sources

Citation linking extends the path for source collection



ProQuest Access Provided by The University of Michigan

ProQuest Dissertations & Theses Global

Enter your search terms... Advanced

Speech Motor Complexity in Childhood Apraxia of Speech

New York University. ProQuest Dissertations Publishing, 2019. 13815309.

Linked References

Full Text - PDF

Abstract/Details

23 References

2 Cited by

References (205)

Model-driven treatment of childhood apraxia of speech: Positive effects of the speech motor learning approach

Children's consonant acquisition in 27 languages: A cross-linguistic review

Changes in movement transitions across a practice period in childhood apraxia of speech

Evaluating significance in linear mixed-effects models in R

Using the Syllable Repetition Task to reveal underlying speech processes in childhood apraxia of speech: A tutorial

Articulatory Control in Childhood Apraxia of Speech in a Novel Word-Learning Task

Intensive Treatment with Ultrasound Visual Feedback for Speech Sound Errors in Childhood Apraxia

Direction of Attentional Focus in Biofeedback Treatment for /R/ Misarticulation

Foundational Research

Foundational Research on this topic

The Treatment Effects of Frequency of Speech Therapy in Early Intervention: A Program Evaluation
Pepe, Danielle. Yeshiva University, ProQuest Dissertations Publishing, 2011

Similar Dissertations

Motor Learning Guided Treatment with Childhood Apraxia of Speech: Cueing & Feedback
Williamson, Sarah. East Carolina University, 2011.

0.91 Similarity index

253 Shared references

Cited by

Foundational Research on this topic

Brief Report: Ages of Language Milestones as Predictors of Developmental Trajectories in Young Childre...
Kover, Sara T; Ed
Ellis Weismer, Su
Autism and Developmer
Disorders, 2016.

Cited by

Cited by 83

Roger; Cox, Martha; et al. Infant and Child Development, 2012.

Assessing Recall in Mothers' Retrospective Reports: Concerns Over Children's Speech and Language Devel...
Russell, Ginny; Miller, Laura L; Ford, Tamsin; Golding, Jean. Journal of Abnormal Child Psychology, 2014

Encoding, memory, and

Similar Dissertations

Similar Dissertations

Motor Learning Guided Treatment with Childhood Apraxia of Speech: Cueing & Feedback
Williamson, Sarah. East Carolina University, 2011

Environmental influences on communication development: Implications for children w...
Aparicio Betancourt, Mariana. University of Illinois at Urbana-Champaign, 2019

Perceptual and acoustical comparisons of motor speech practice options for children with childhood apraxia of sp...
Nordness, Amy S. The University of Nebraska - Lincoln, 2011

The Test of Tim Speech Timing Differentiate Ch Without Speech
Tone, Tori N. 2020

ProQuest ExLibris Part of Clarivate

A new partnership with Research4Life

in alignment with Clarivate’s sustainability mission and United Nation SDGs

SUSTAINABLE DEVELOPMENT GOALS

Donating FREE ACCESS to institutions in 95 lower-income countries--part of

research4life

https://www.research4life.org/other/95-research4life-countries-granted-pqdt-access/

“Our goal is to help close knowledge gaps, unify research, and elevate the scholarship being conducted in every corner of the world.”

- Angela D’Agostino, VP Dissertations & Theses,
ProQuest part of Clarivate

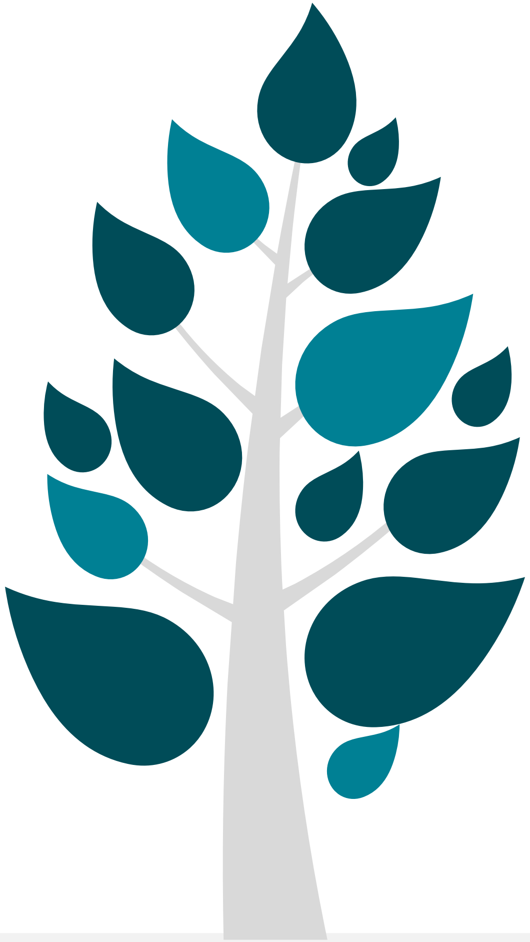
The Vision: Harnessing the Power of Data-driven Research Insights

ProQuest
Dissertations & Theses
GLOBAL

+

Clarivate
Web of Science™

Dissertations and Theses are essential for comprehensive discovery



Web of Science™

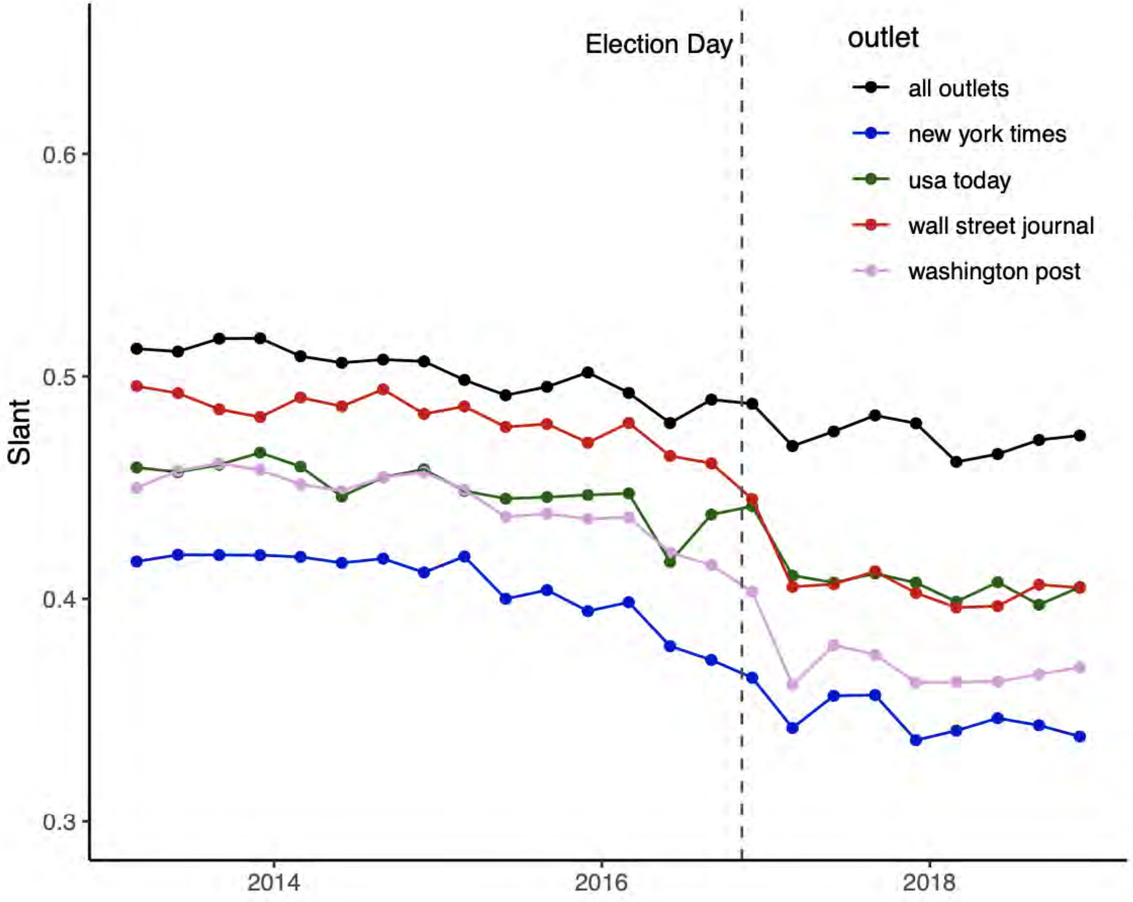
- Peer-reviewed Journal Articles
- Conference Proceedings
- Books
- Emerging Sources

ProQuest Dissertations & Theses Global

- Unpublished scholarship
- Unpublished data sets
- Marginalized perspectives

Journalist Ideology and the Production of News: Evidence from Movers (joint with Jacob Conway)

- Train RoBERTa-based model to measure slant using news articles tweeted by US politicians.
  - Significant improvement over prior bag-of-word methods
- Apply ML model to 20+ million full-text articles in TDM dataset.
- Document novel descriptive evidence on US media slant over time.
- Measure how slant changes as journalists move between outlets (journalist preferences explain 16% of variation).



LINK TO THE PAPER: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4295587

Abstract(s) at the core: A case study of disciplinary identity in the field of linguistics

Taylor LiCausi & Daniel McFarland

- Leveraged the ProQuest Thesis ID variable to construct the broader field of linguistics.
- Used dissertation abstracts from 1980-2010 to create a structural topic model (STM).
- Mapped trends of our 40 topics by degree type and calculated topic correlations.

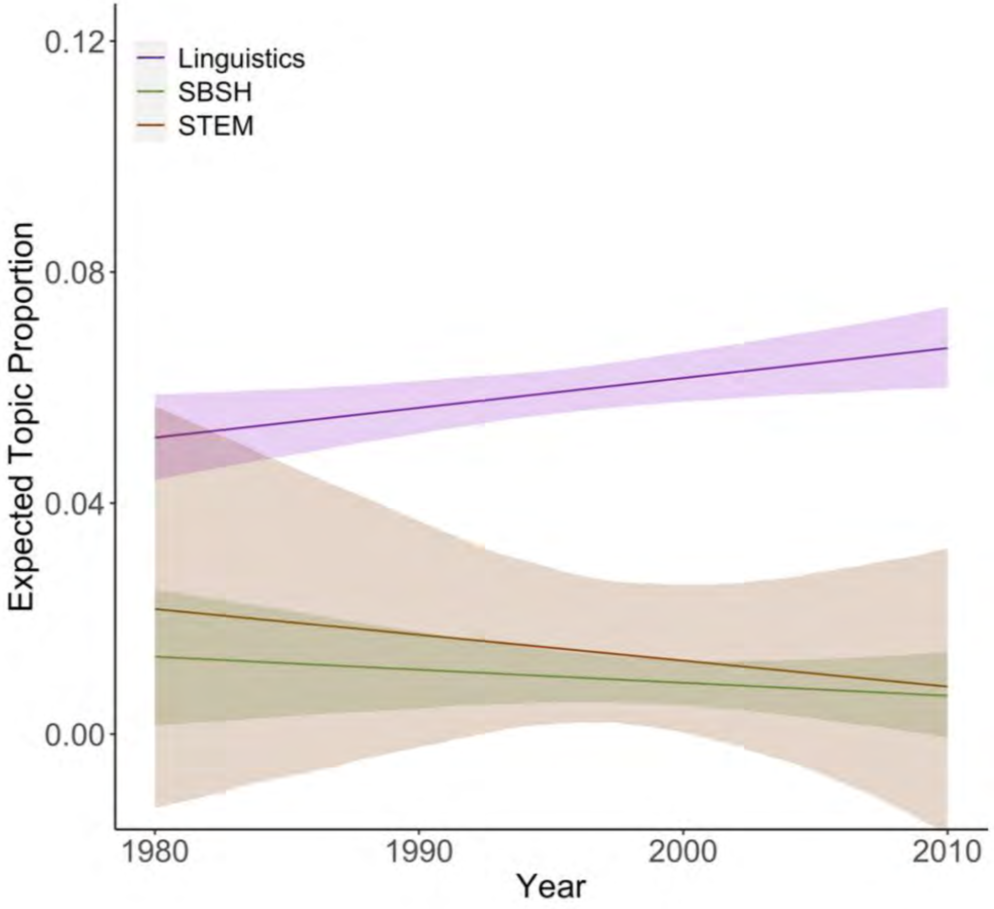
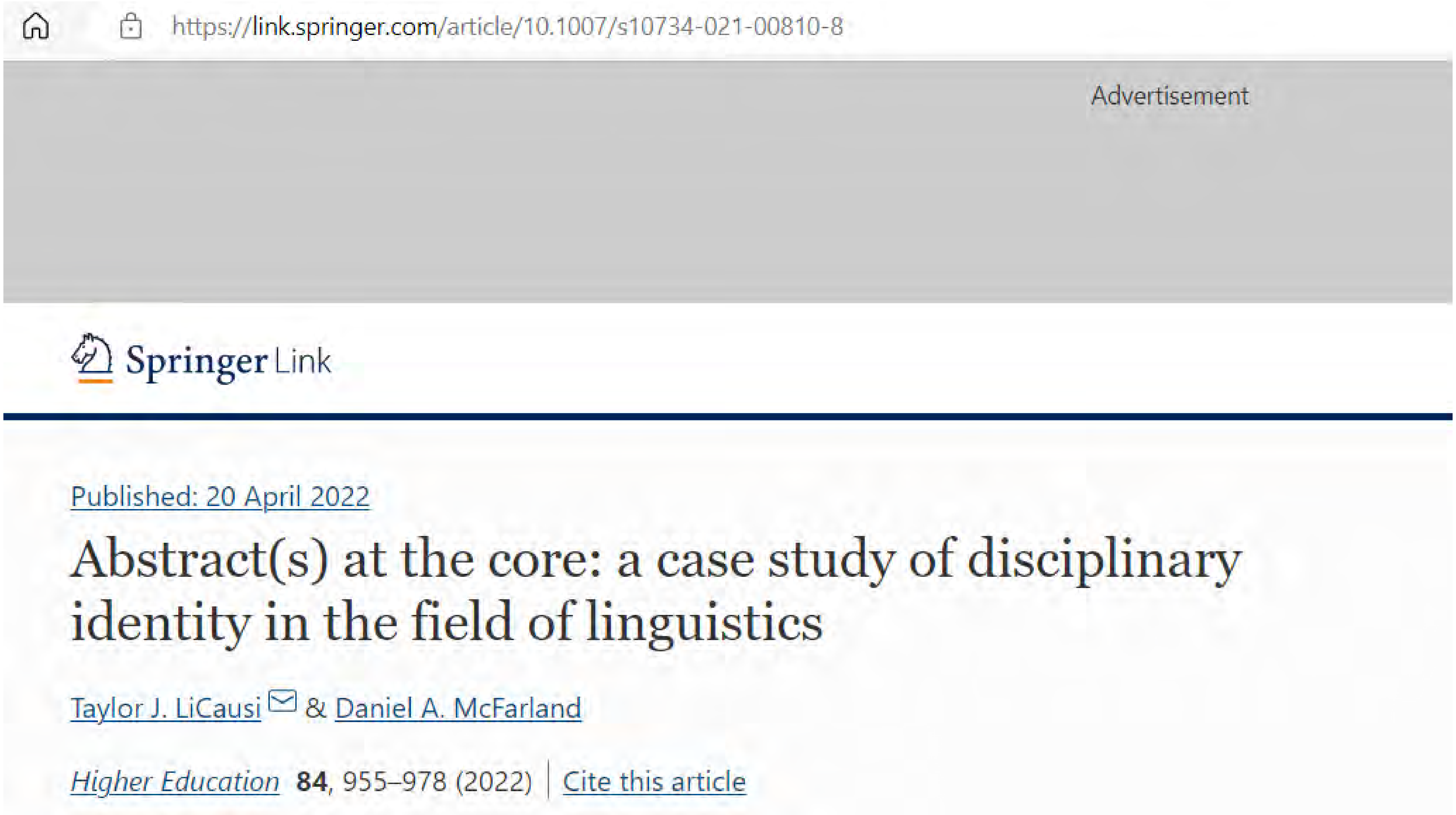


Fig. 1 Expected topic proportion of syntax: main (8), 1980-2010

Results

- Sketched an intellectual topography of the broader field of linguistics
- Identified linguistics’ disciplinary core through the topics most distinctive to linguistics degree abstracts.
- Found that STEM disciplines were more likely to engage with linguistics’ core than SBSH disciplines.
- Found evidence of differentiation for SBSH and STEM topics.

Link to LiCausi & McFarland’s article: Abstract(s) at the core: a case study of disciplinary identity in the field of linguistics | SpringerLink



https://link.springer.com/article/10.1007/s10734-021-00810-8

Published: 20 April 2022

Abstract(s) at the core: a case study of disciplinary identity in the field of linguistics

Taylor J. LiCausi & Daniel A. McFarland

Higher Education **84**, 955–978 (2022) | Cite this article

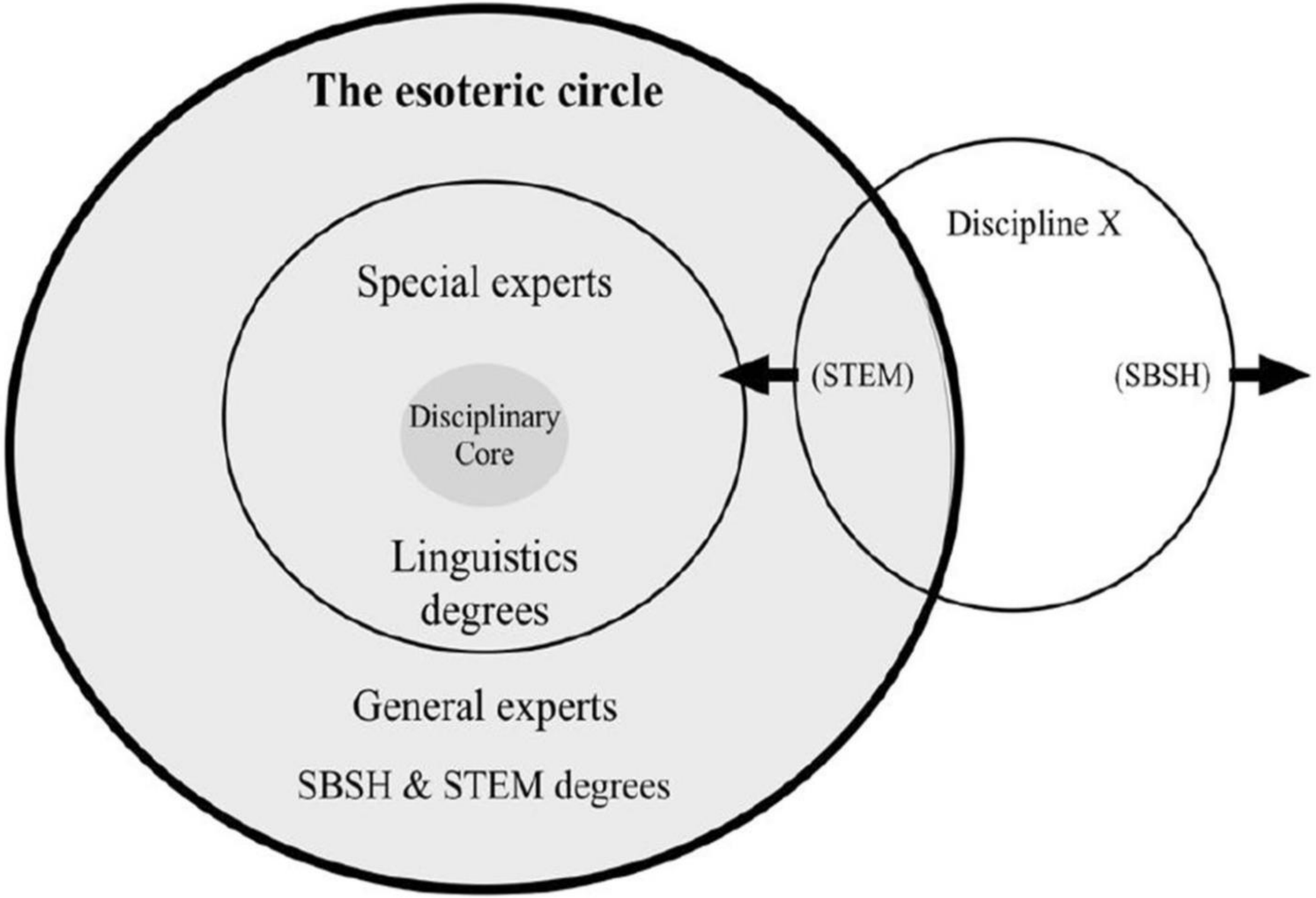


Fig. 13 The esoteric circle of linguistics

ProQuest
TDM Studio

SIGN UP FOR A FREE 30-Day TRIAL

https://about.proquest.com/en/products-services/TDM-Studio/

ProQuest TDM Studio

allows you to…

Access and Analyze

Leverage content you already subscribe to and save months of time and tens of thousands of dollars required to negotiate and acquire data mining licensing. TDM Studio provides access to content with text and data mining rights already cleared.

Target the Content

Streamline the content selection process. Refine and target the dataset to your research interest, eliminating the need to process unrelated documents.

Connect and Collaborate

Enables real time collaboration across multiple sites and researchers in one workbench.

ProQuest
TDM Studio

CLEARED LICENSES FOR TDM

- Over 3K newspapers
- Dissertations & Theses
- Primary Sources
- Government documents (in BETA)

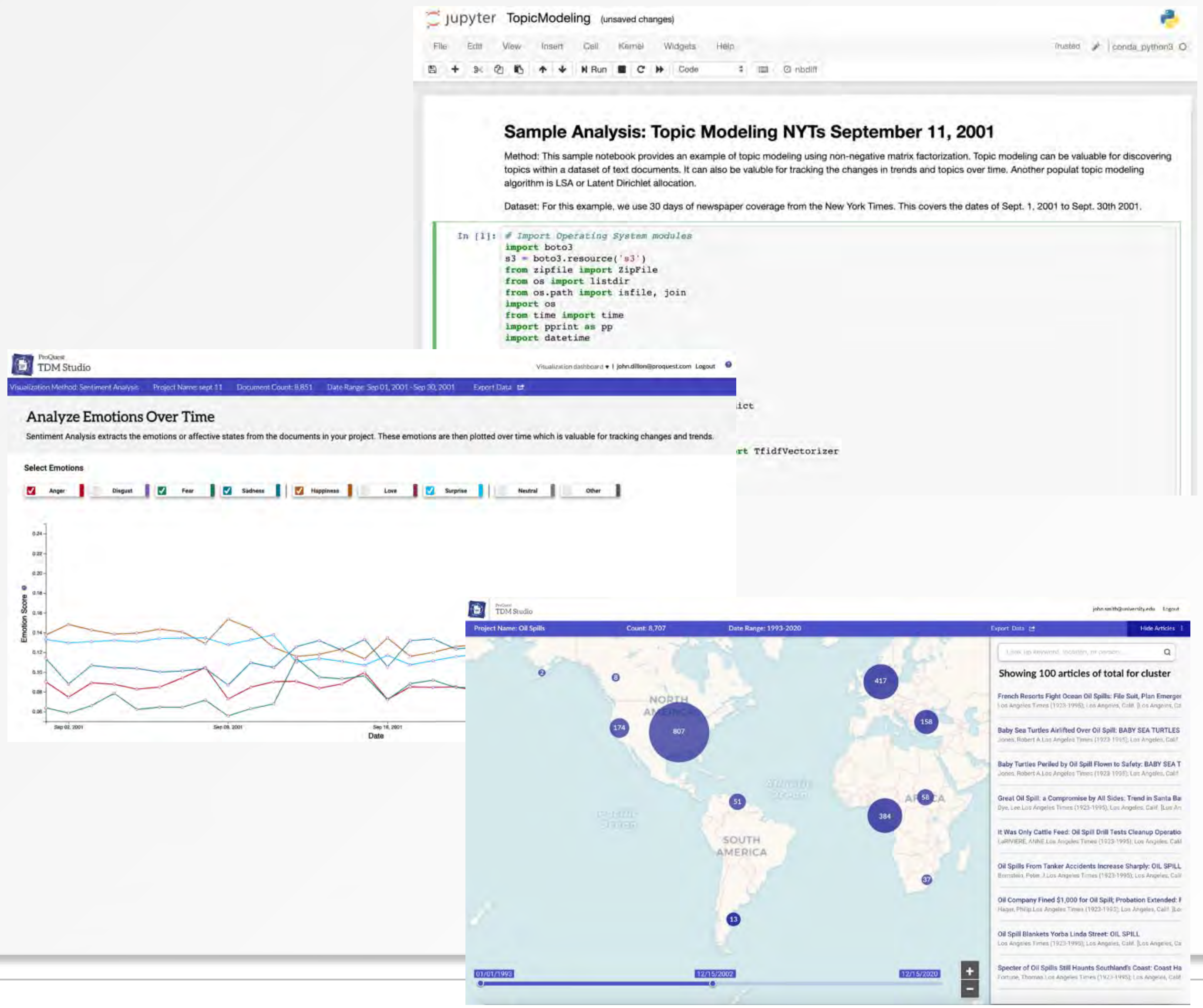
WORKBENCHES for R and Python

VISUALIZATION TOOLS

- Sentiment Analysis
- Topic Modeling
- Geographic Analysis

MOST OFTEN USED FOR

- Information Science
- Social Sciences
- Economics
- Political Science
- Business



Sample Analysis: Topic Modeling NYTs September 11, 2001

Method: This sample notebook provides an example of topic modeling using non-negative matrix factorization. Topic modeling can be valuable for discovering topics within a dataset of text documents. It can also be valuble for tracking the changes in trends and topics over time. Another populat topic modeling algorithm is LSA or Latent Dirichlet allocation.

Dataset: For this example, we use 30 days of newspaper coverage from the New York Times. This covers the dates of Sept. 1, 2001 to Sept. 30th 2001.

```
In [1]: # Import Operating System modules
        import boto3
        s3 = boto3.resource('s3')
        from zipfile import ZipFile
        from os import listdir
        from os.path import isfile, join
        import os
        from time import time
        import pprint as pp
        import datetime
ict
rt TfidfVectorizer
```

Visualization Method: Sentiment Analysis | Project Name: sept 11 | Document Count: 8,851 | Date Range: Sep 01, 2001 - Sep 30, 2001 | Export Data

Analyze Emotions Over Time

Sentiment Analysis extracts the emotions or affective states from the documents in your project. These emotions are then plotted over time which is valuable for tracking changes and trends.

Select Emotions

☑ Anger □ Disgust ☑ Fear ☑ Sadness ☑ Happiness □ Love ☑ Surprise □ Neutral □ Other

Project Name: Oil Spills | Count: 8,707 | Date Range: 1993-2020 | Export Data | Hide Articles

Showing 100 articles of total for cluster

French Resorts Fight Ocean Oil Spills: File Suit, Plan Emerge
Los Angeles Times (1923-1995); Los Angeles, Calif. [Los Angeles, Ca

Baby Sea Turtles Airlifted Over Oil Spill: BABY SEA TURTLES
Jones, Robert A.Los Angeles Times (1923-1995); Los Angeles, Calif

Baby Turtles Periled by Oil Spill Flown to Safety: BABY SEA T
Jones, Robert A.Los Angeles Times (1923-1995); Los Angeles, Calif

Great Oil Spill: a Compromise by All Sides: Trend in Santa Ba
Dye, Lee.Los Angeles Times (1923-1995); Los Angeles, Calif. [Los An

It Was Only Cattle Feed: Oil Spill Drill Tests Cleanup Operatio
LaRIVIERE, ANNE.Los Angeles Times (1923-1995); Los Angeles, Calif

Oil Spills From Tanker Accidents Increase Sharply: OIL SPILL
Bornstein, Peter J.Los Angeles Times (1923-1995); Los Angeles, Calif

Oil Company Fined $1,000 for Oil Spill; Probation Extended: F
Hager, Philip.Los Angeles Times (1923-1995); Los Angeles, Calif. [Lo

Oil Spill Blankets Yorba Linda Street: OIL SPILL
Los Angeles Times (1923-1995); Los Angeles, Calif. [Los Angeles, Ca

Specter of Oil Spills Still Haunts Southland's Coast: Coast Ha
Fortune, Thomas.Los Angeles Times (1923-1995); Los Angeles, Calif

Thanks for joining us!

gilia.smith@clarivate.com

ProQuest

Part of Clarivate