

Introduction

We present an educational tool that integrates computational linguistics resources for use in non-technical undergraduate language science courses. Using the tool with evidence-driven pedagogical case studies, we provide opportunities for students to gain an understanding of linguistic concepts and analysis through the lens of realistic problems.

Case Studies

We developed three case studies for a course on language science fundamentals and an English language history course. Combining case studies with the functionality of Linguine provides fictional yet realistic scenarios to solve using active inquiry with linguistic data. Each case study has:

- A **narrative** that sets up a problem and background
- **Data** prepared for in- and out-of-class analysis
- Step-by-step **analysis instructions** with **questions** to promote critical thinking
- Evaluation **rubric** for oral and written reporting and a quiz
- **Readings** to consult: academic; applied broad-audience

The Language of Dementia

Students analyze a selection of picture descriptions from the DementiaBank¹ corpus with the goal of assisting a medical researcher in identifying linguistic markers of Alzheimer's disease, using features such as disfluencies, lexical complexity, n-grams, and syntactic tree structures.

Historical Varieties of English

Students examine excerpts of literature across time periods to assist school teachers in choosing grade-appropriate readings for their classes using characteristics such as text readability, syllable counts, and parts-of-speech.

Formality in Business Communications

In roles as analysts for a training agency, students use email data to critically envision guidelines for professional workplace communications, exploring features such as named entities, sentiment, and syntactic tree structures.

Evaluating the Approach

Q: This case study activity...

- 1 was engaging.
- 2 had clear instructions.
- 3 was related to the course material.
- 4 involved a reasonable time commitment.
- 5 was a valuable learning experience.
- 6 enhanced my understanding of linguistic concepts.
- 7 reinforced theoretical concepts from with an application.
- 8 let me use linguistic approaches to problem solving.
- 9 had me engaged in critical thinking.
- 10 involved a useful reporting experience.

Q: Using the provided web tools and input...

- 11 was straightforward.
- 12 went hand-in-hand with the case instructions.
- 13 enhanced my thinking about the case resolution plan.
- 14 was interesting.
- 15 was a good learning experience.

Table 1: Case study satisfaction survey agreement statements.

Linguine

- Offers a straightforward and intuitive interface.
- Enables class activities and active learning.
- Allows interaction with preloaded or custom-uploaded plain texts and their visualizations.
- Transforms machine-processed results into intuitive visualizations.

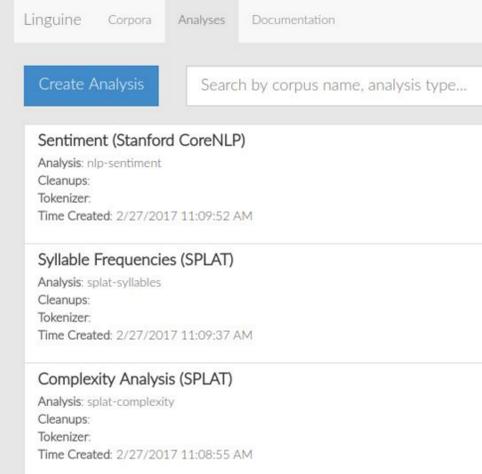


Fig. 1: Linguine's tab-based interface: Analyses tab with 3 completed analyses.

Unigram Frequencies	Bigram Frequencies	Trigram Frequencies
11 a	1 a black	1 a black horse
2 about	1 a blue	1 a blue coat
2 accept	1 a certain	1 a certain step
1 admitted	1 a cousin	1 a cousin and
1 advantage	1 a few	1 a few days
1 after	2 a large	2 a large party
1 afterwards	1 a little	1 a little by
1 agreeable	1 a number	1 a number of
4 all	1 a report	1 a report soon
1 already	1 a sight	1 a sight of
1 altogether	1 about from	1 about from one
1 always	1 about ten	1 about ten minutes

Fig. 2: N-gram frequencies for a sample of Jane Austen's writing.

1st-Person 2nd-Person 3rd-Person

Pronoun	Frequency	Type
I	1	Personal, Singular
HER	2	Personal/Possessive, Singular/Plural

Fig. 3: Pronoun frequencies for the sentence *I saw her walking her dog on Monday.*

Sentence Sentiment: Very positive

This is the best news I've heard all year !

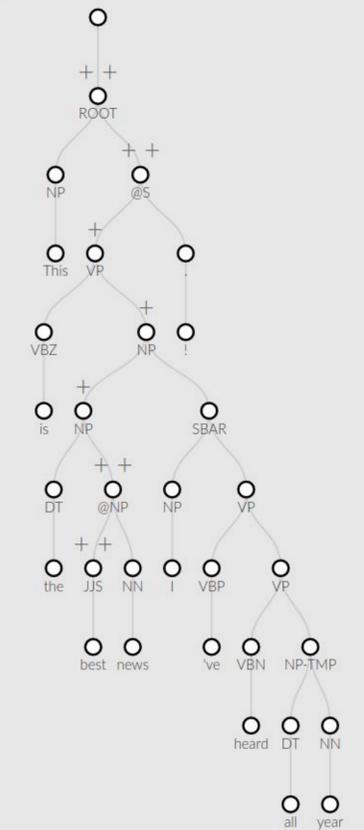


Fig. 4: Syntactic tree with sentiment labels for the sentence *This is the best news I've heard all year!*

Complexity Metric	Value
Content Density:	1.206
Idea Density:	0.499
Flesch Readability:	72.000
Flesch-Kincaid Grade Level:	10.700
Type-Token Ratio:	0.469 (283 / 603)

Fig. 5: Complexity metrics for an excerpt of Sir Arthur Conan Doyle's writing.

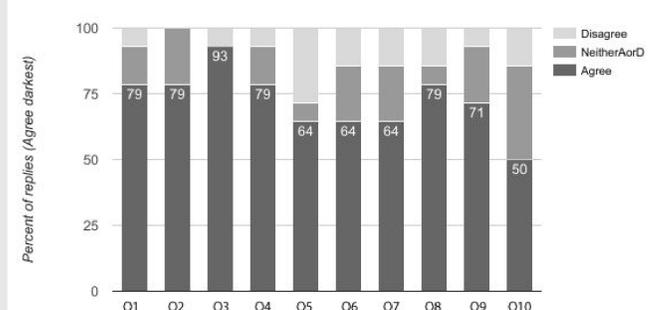
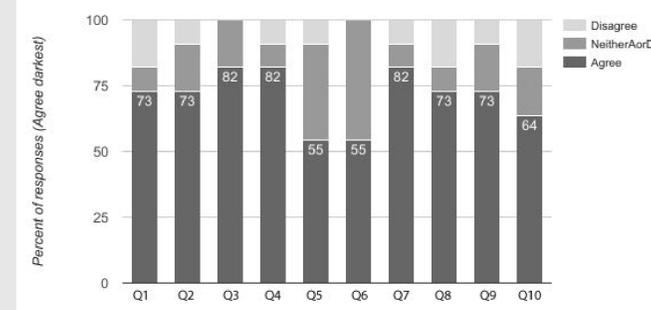


Fig. 6: Agreement ratings from students to the questions in Table 1 for a case study in the introductory language science course.



Instructor Observations

- Students exhibited increased critical thinking about analysis, data, and methods.
- Reporting exercises nurtured co-learning: students observed how others approach a problem and develop different analyses and visualizations.
- Cases offered a structured framework for teamwork.

Student Feedback

- Majority of students evaluated the activity positively, agreeing that the experience was engaging, educational, interesting, and stimulated critical thinking and learning.
- Students reported that the activity and tool were clear and straightforward to use.
- Most students felt the experience was relevant to and practiced class material, enhanced their understanding of linguistics, and engaged them in problem solving.

I learned that there are real world applications for what we learned in class.

I was able to come up with something that could prove to be useful in my life

I really enjoyed using linguine, it's a great tool.

got to work with real data on solving a real problem

Interesting scenario, engaging

n-grams could be downloadable as a csv for additional processing

I wish there was a way to get quantitative data about the parse trees

Table 2: Open-ended student feedback responses.

Conclusions

- Integration of the pedagogical web application, Linguine, with case studies is an effective learning approach for language science students.
- Possible future work:
 - Improving Linguine's performance and effectiveness
 - Expanding and creating new case studies
 - Designing a systematic process for preparing new case study materials
 - Improvement of existing visualizations
 - Expanding analysis potential for transcribed speech

Acknowledgments

This work was partially supported by an RIT PLIG grant. Many thanks to the three senior project teams that helped develop Linguine's core functionality and interface: ProNouns, Rigatoni, and Pastafarians.

References

- ¹ J.T. Becker, F. Boiler, O.L. Lopez, J. Saxton, and K.L. McGonigle. 1994. The natural history of Alzheimer's disease: Description of study cohort and accuracy of diagnosis. *Archives of Neurology* 51(6):585-594.
- ² J. Goldstein, M. Starbard, and S. Wheeler (Producers). 2010. *Agatha Christie and nuns tell a tale of Alzheimer's*. NPR [June 1].
- ³ G. Szatloczki, I. Hoffmann, V. Vincze, J. Kalman, and M. Pakaski. 2015. Speaking in Alzheimer's disease, is that an early sign? Importance of changes in language abilities in Alzheimer's disease. *Frontiers in Aging Neuroscience* 7:195.
- ⁴ Katharine Perera. 1980. The assessment of linguistic difficulty in reading material. *Educational Review* 32(2):151-161.
- ⁵ Scott Paris. N.d. *Understanding text complexity*. McGraw-Hill Education.
- ⁶ Elie Pavlick and Joel Tetreault. 2016. An empirical analysis of formality in online communication. *TACL* 4:61-74.
- ⁷ Shana Lebowitz. 2015. *Here's exactly how to write an email to your CEO*. *Business Insider* [Nov. 13].