ESL Article Acquisition

Elena Cimino
April 11, 2019

Outline

- Background information
  - Article research
- Motivations and research questions
- My project
  - Explanation of corpoa
- Descriptive stats
- Data samples and preliminary analysis

Background: English article system

- English article system difficult for non-native speakers to acquire
  - Definite *the*, indefinite *a/an*, zero article $\emptyset$
- Zhao and MacWhinney (2018): articles and cues in English
  - The forms of English articles map onto approximately 90 different functions
  - Article cues indicate what form should be used

- Two types of article cues:
  - *General:* easy to generalize
    - e.g. construction names → *the*
    - I took a picture of the Hot Metal Bridge.
  - *Idiosyncratic:* based on a smaller lexical field
    - e.g. exceptions to construction names → Ø
    - I took a picture of Ø Alumni Hall.

- Much of the literature suggests an L1 effect (e.g., Butler, 2002; Ionin et al., 2008; Master, 1997; Stehle, 2009; Zobl, 1982):
  - Speakers with [+art] L1 tend to overgeneralize *the* and underutilize *a/an*
  - Speakers with [-art] L1 initially overgeneralize usage of Ø, eventually move toward *the* as default
    - Have higher accuracy with *a/an*
  - [+art] L1 speakers tend to move more quickly toward native-like article usage than [-art] L1 speakers

Background: article research

- In this area of research, focus tends to be on comparing:
  - "Full" [+art] languages (e.g. Spanish, German)
  - [-art] languages (e.g. Chinese, Korean, Russian)

- Arabic is an interesting cross-section
  - "Partial" [+art] with definite article "al-" (الـ)
  - No indefinite article
  - Rarely represented in this literature

How do L1 speakers of Arabic acquire English articles with respect to speakers of L1s that are "fully" [+art] and speakers of L1s that are [-art]?

Research questions

1. Is there an L1 effect in ESL article acquisition?
   - How does Arabic pattern?

2. Which articles do L2 learners of English have the most difficulty acquiring?
   - Are there specific functions that are difficult?
   - Are there specific nominal features that are particularly troublesome?

◆ Comparing 2 corpora of L2 English learners' writing
- ◆ BALC – BuiD Arabic Learner Corpus
  - ◇ Arabic L1 only, compiled in United Arab Emirates
  - ◇ Targets: essay responses to the Common Education Proficiency Assessment (CEPA)
- ◆ PELIC – Pitt ELI Corpus
  - ◇ Many L1s, compiled in Pittsburgh
  - ◇ Targets: Arabic, Korean, and Spanish L1 learners

My project

BALC

- 1,865 texts total
  - End up with 1,676 TXT files (CEPA only)
  - 6 Levels
    - 1: random
    - 6: may be ready for direct entry

PELIC

Number of Texts per Level (BALC)
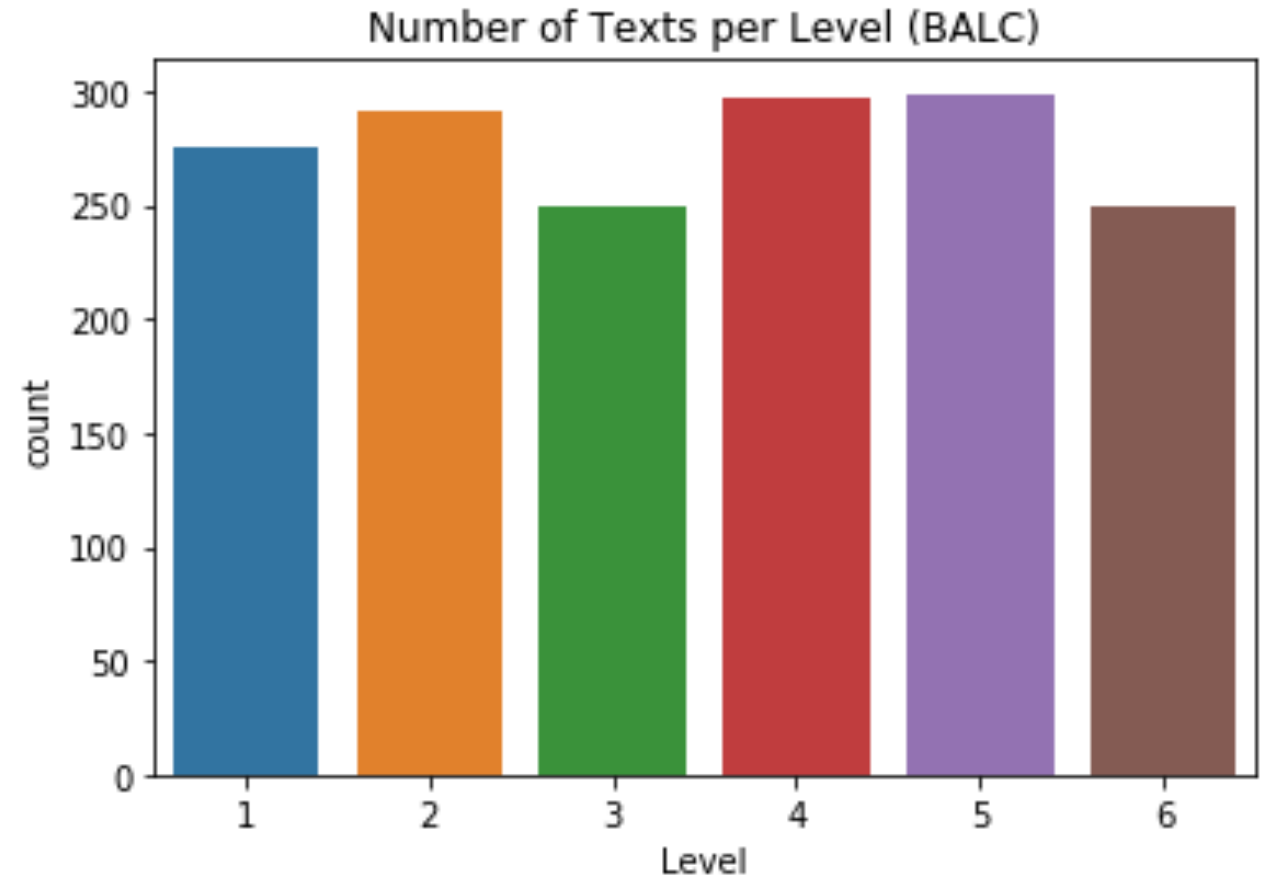
My project

BALC

- 1,865 texts total
  - End up with 1,676 TXT files (CEPA only)
  - 6 Levels
    - ◇ 1: random
    - ◇ 6: may be ready for direct entry

PELIC

- Written:
  - 48,000+ texts
  - After culling, 5,052 usable texts
- Levels 1 → 5
  - 1: High beginning
  - 5: Low Advanced
- Also has spoken data (unused for this project)

CEPA-English

- Used as a measure of students' English ability in UAE
  - Originally used as a placement test (c. 2003), later became an admissions test

- Has three parts:
  - ◇ Grammar and Vocabulary
  - ◇ Reading
  - ◇ Writing
    - Prompt in both English and Arabic
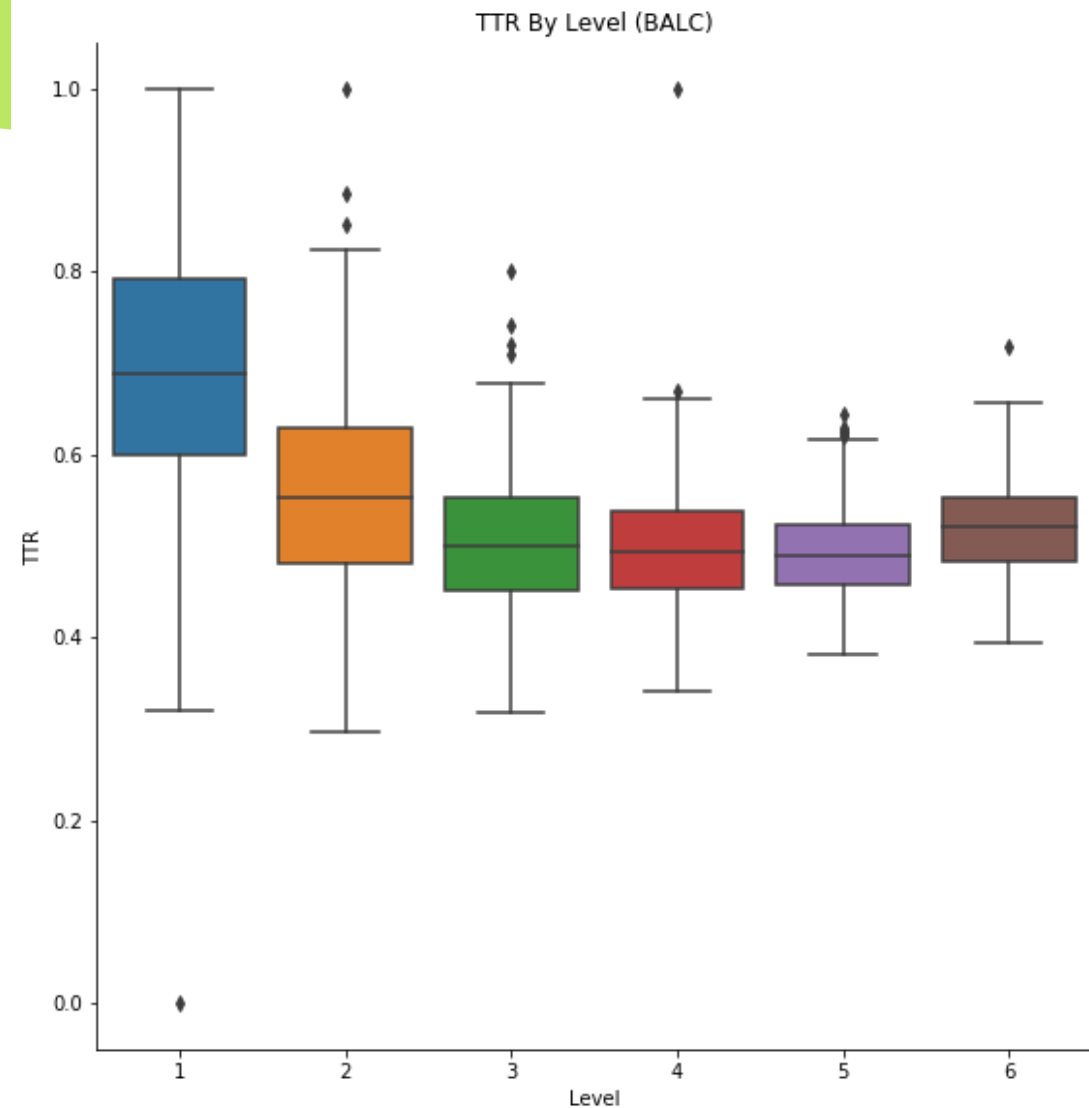   - 30 minutes to write a 150-200 word essay by hand

The fun part!

Some basic stats (BALC)

Token Count by Level (BALC)

TTR By Level (BALC)
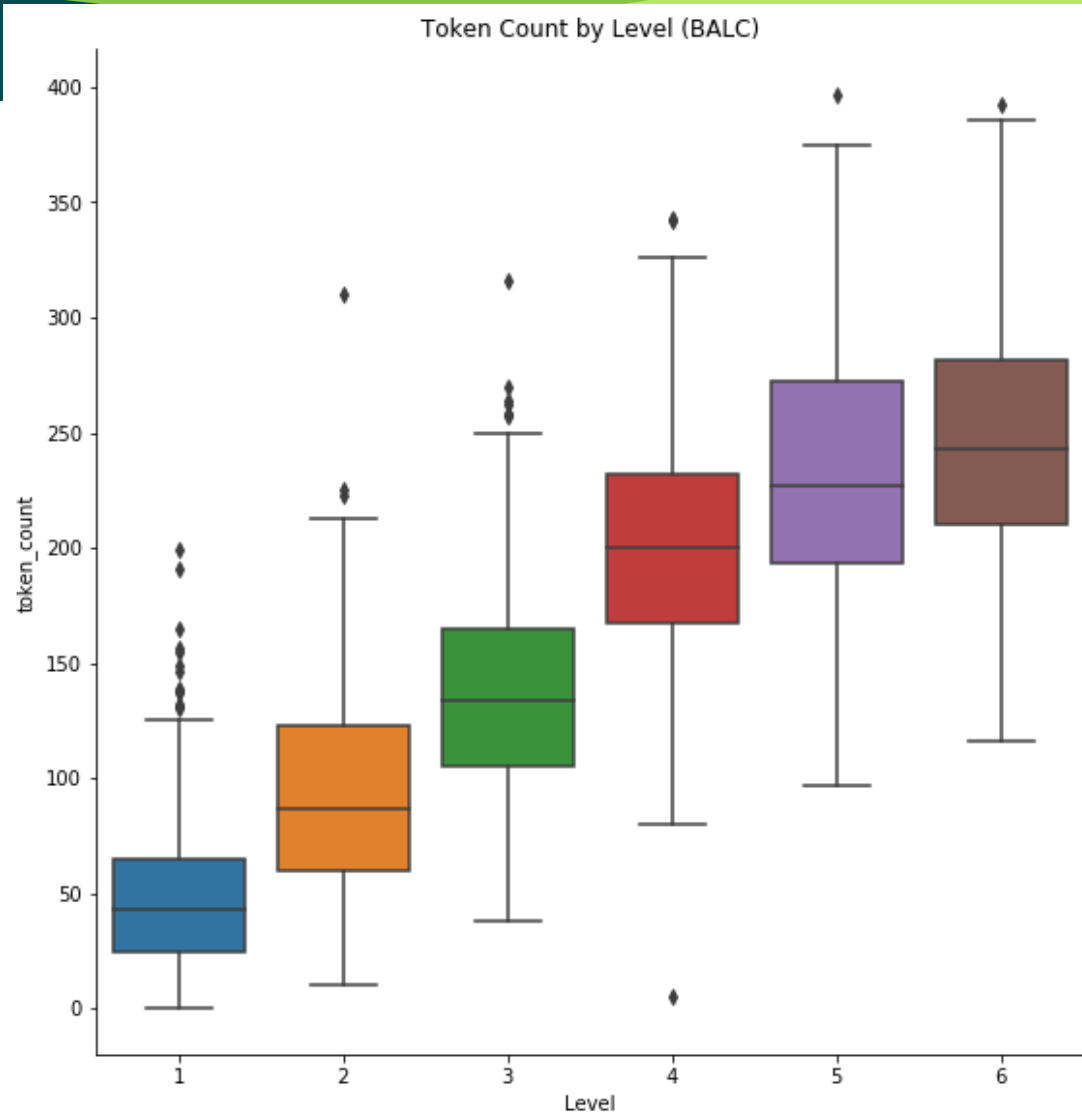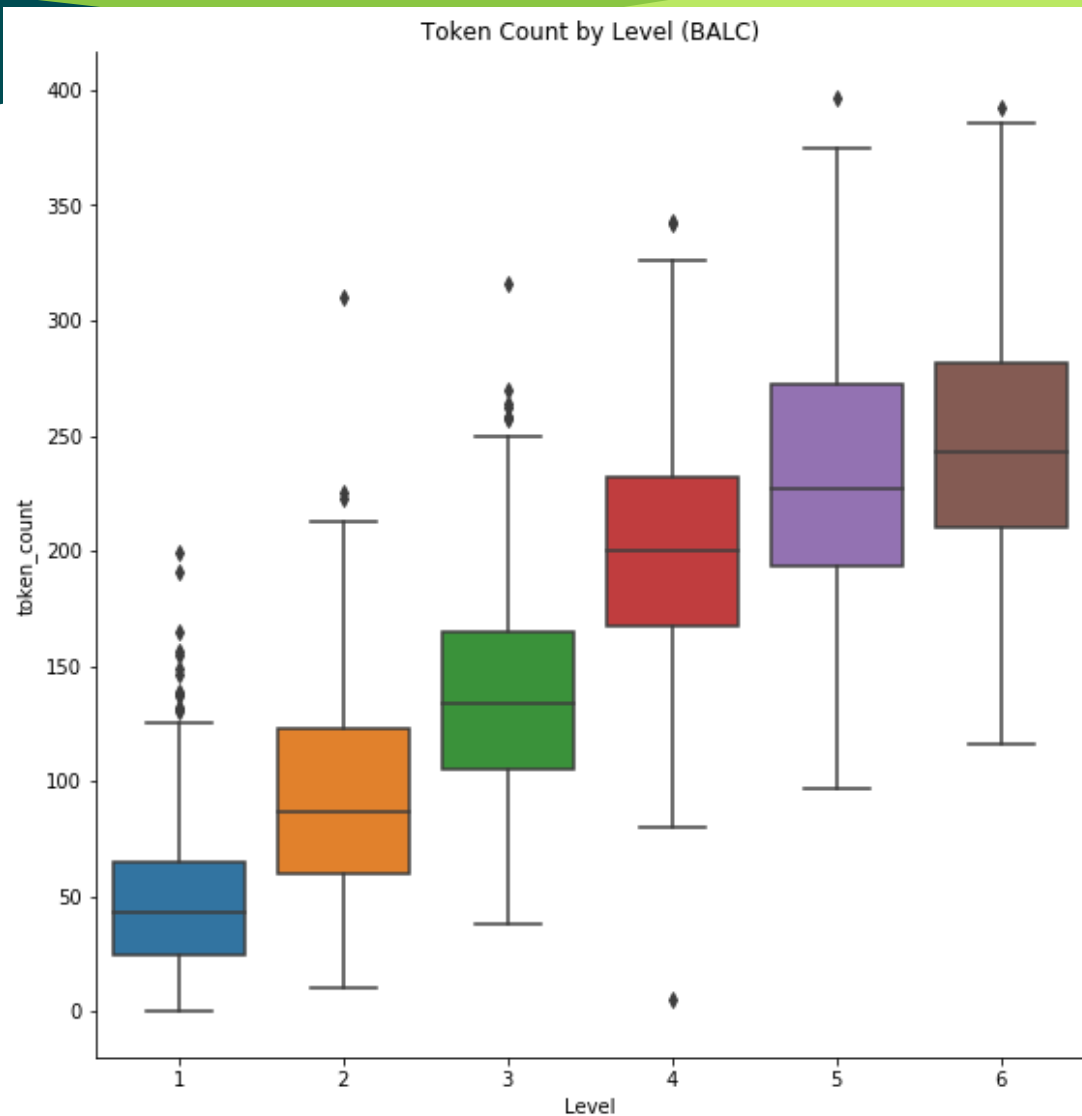
F_onewayResult(statistic=764.373757392868, pvalue=0.0)

F_onewayResult(statistic=208.1989467995349, pvalue=1.603e-172)

Token Count by Level (BALC)

Guiraud's R by Level (BALC)

$$Guiraud's\ R = \frac{Types}{\sqrt{Tokens}}$$

15

F_onewayResult(statistic=764.373757392868, pvalue=0.0)
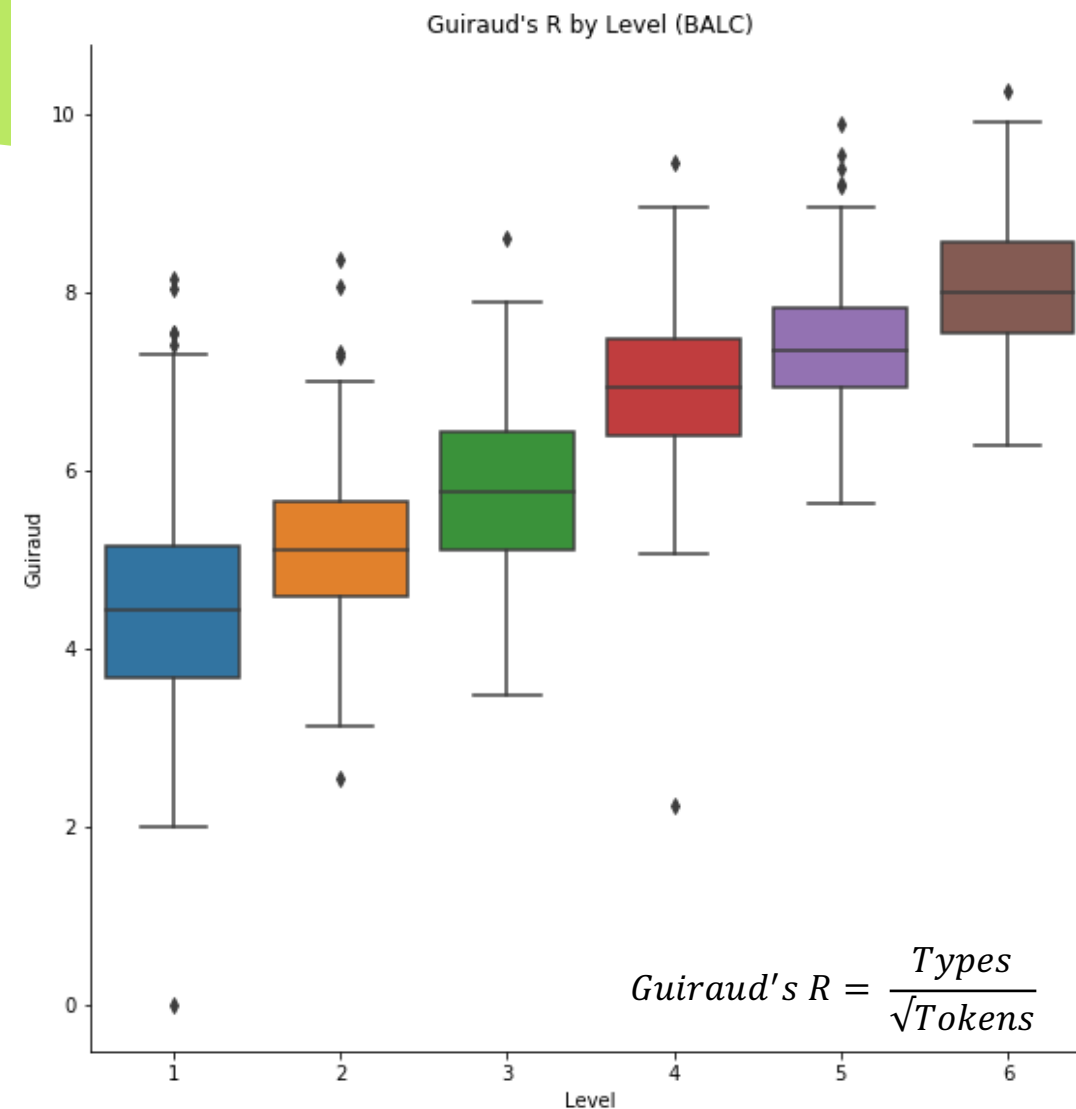
F_onewayResult(statistic=716.2910170871269, pvalue=0.0)

Some basic stats (PELIC)

◆ Dropped levels 1 and 2

◆ Capped token length at 400 (to stay comparable with BALC essays)

◆ Used only texts from Writing class

Some basic stats (PELIC)

- ◆ Dropped levels 1 and 2
- ◆ Capped token length at 400 (to stay o[...] essays)
- ◆ Used only texts from Writing class
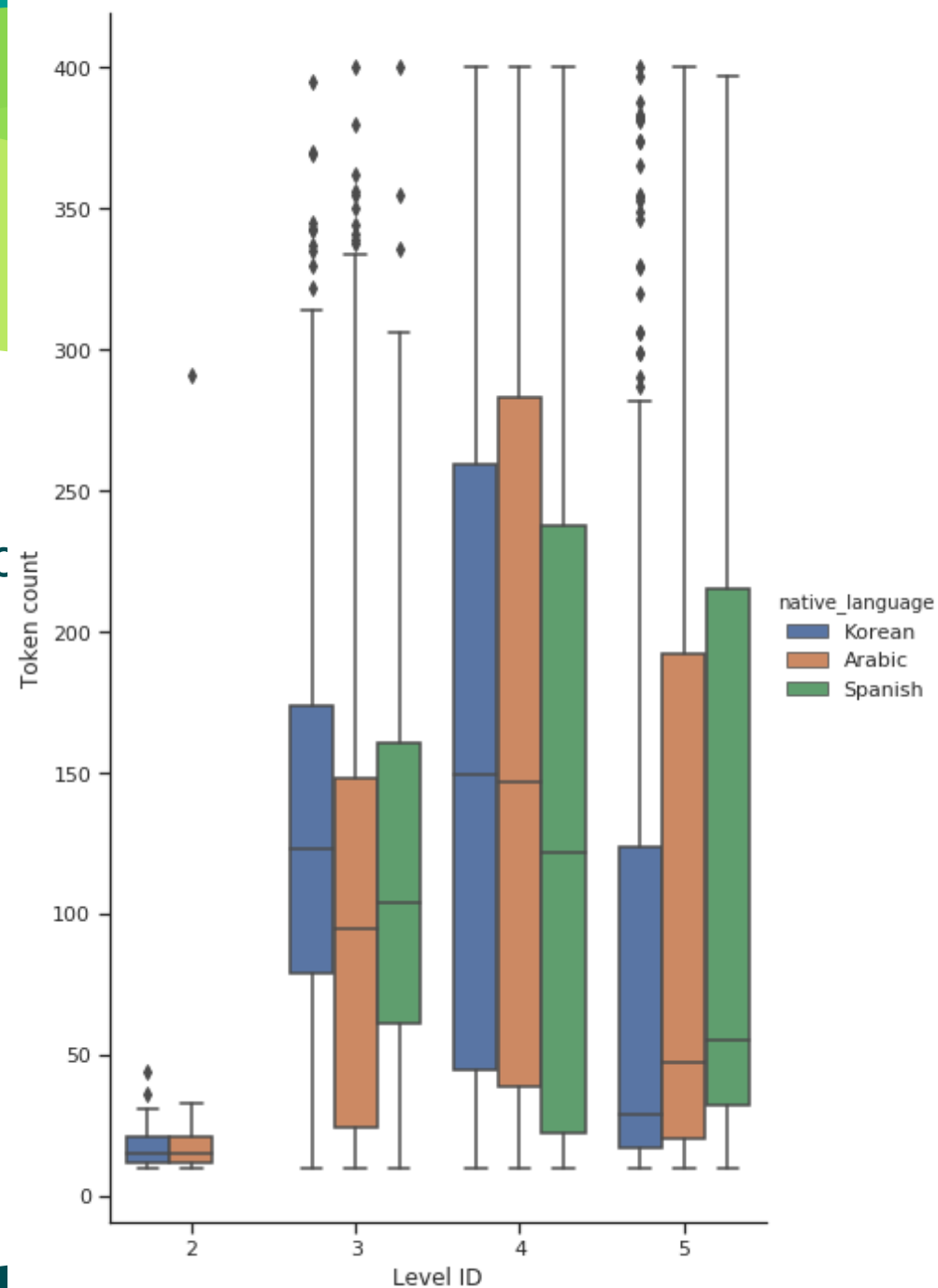
Picking samples

- Target specific nouns:
  - 5 count (place, park, week, story, book)
  - 5 non-count (time, sea, food, water, money)
  - Not a perfect selection, but some logistical constraints
- Will analyze nominal objects and articles in essays as a whole

Samples

BALC

Level	count	mean	std	min	25%	50%	75%	max
3	10.0	151.7	64.187313	78.0	102.25	144.0	167.00	264.0
4	10.0	211.3	41.411351	138.0	185.75	209.0	238.00	286.0
5	10.0	251.6	70.062195	172.0	188.75	255.0	282.25	397.0
6	10.0	229.4	40.467271	183.0	200.75	220.5	243.00	303.0

Token count

PELIC

level_id	native_language	count	mean	std	min	25%	50%	75%	max
3	Arabic	10.0	173.0	103.825709	10.0	107.75	167.0	208.25	356.0
	Korean	10.0	171.9	58.856605	100.0	127.50	156.5	207.00	280.0
	Spanish	10.0	159.6	78.785503	69.0	116.75	160.0	171.00	355.0
4	Arabic	10.0	205.6	132.624114	13.0	86.75	223.5	309.75	384.0
	Korean	10.0	226.1	98.635637	46.0	169.75	206.5	319.75	356.0
	Spanish	10.0	220.5	134.122871	59.0	98.50	189.0	342.50	398.0
5	Arabic	10.0	258.2	128.273839	20.0	160.25	284.5	370.75	390.0
	Korean	10.0	144.5	103.596707	28.0	68.00	140.0	180.25	374.0
	Spanish	10.0	290.0	113.430351	33.0	270.75	330.0	361.50	397.0

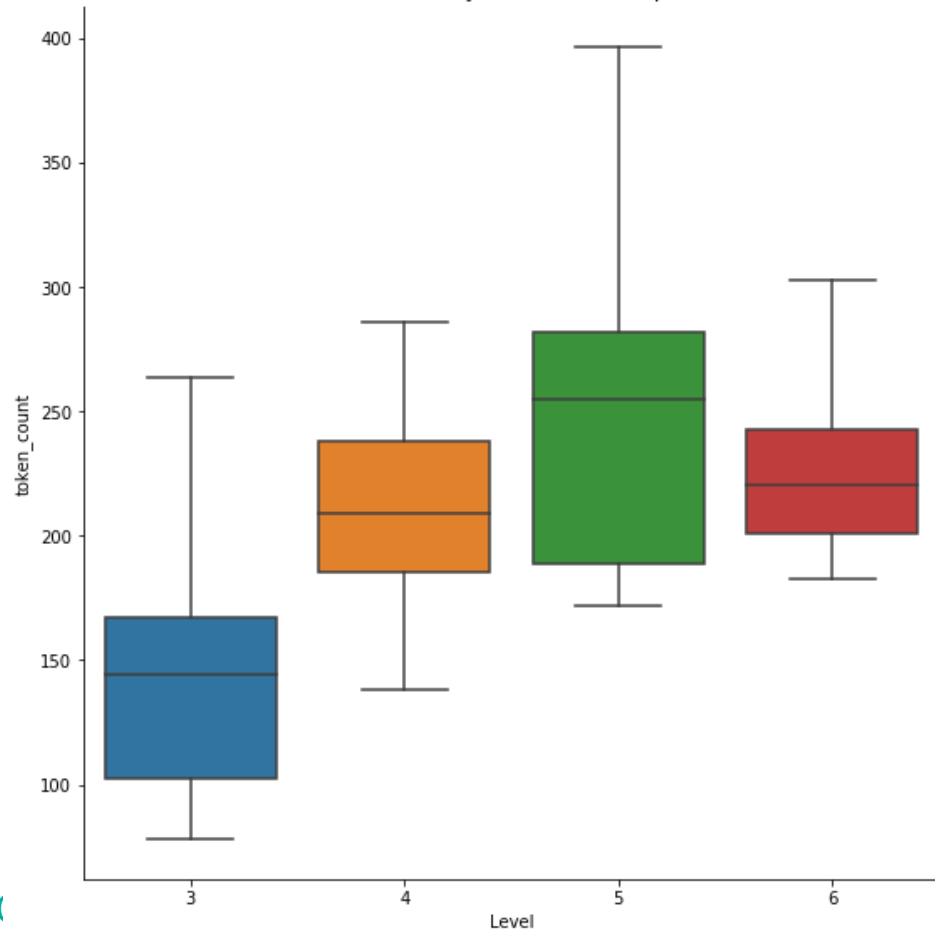# Samples

BALC

Token count

PELIC

Token Count by Level (BALC samples)

Token Count by Level and L1 (PELIC samples)
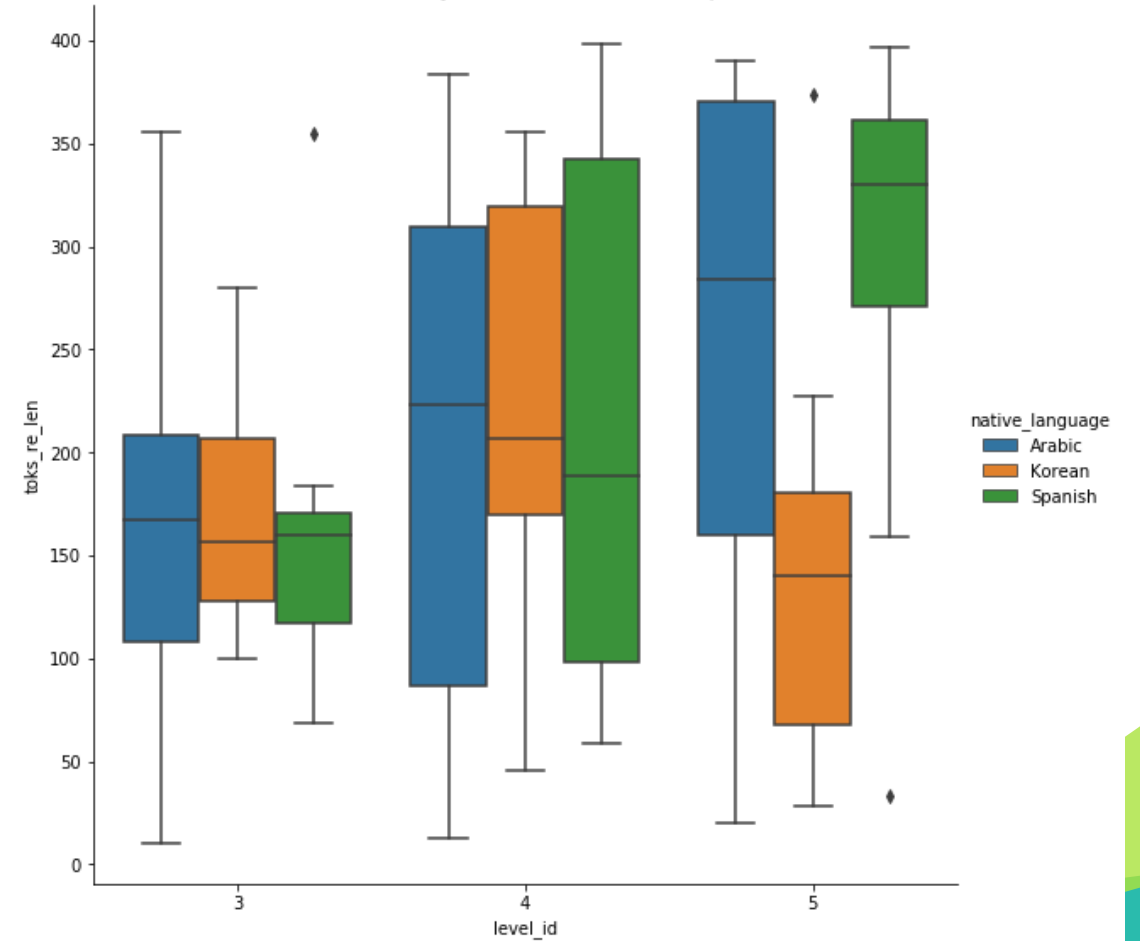
Samples

BALC

Level	count	mean	std	min	25%	50%	75%	max
3	10.0	6.163016	1.109401	4.858987	5.306816	5.996125	7.194436	7.877855
4	10.0	6.811704	1.021188	5.344792	6.127009	6.680361	7.663461	8.455767
5	10.0	7.484633	0.928882	6.186455	6.807033	7.507225	7.826307	8.944468
6	10.0	8.029801	0.465230	7.529412	7.798222	7.981108	8.169718	9.152086

Guiraud's R

PELIC

| level_id	native_language	count	mean	std	min	25%	50%	75%	max
3	Arabic	10.0	6.410718	1.841603	3.162278	4.989269	6.692213	7.795469	9.009982
	Korean	10.0	6.780253	1.681038	3.502303	5.884419	6.769231	7.943260	9.203260
	Spanish	10.0	6.093174	0.596275	5.160468	5.808045	5.964809	6.427172	7.295815
4	Arabic	10.0	7.055018	2.018816	3.605551	5.420555	7.629763	8.064704	10.190197
	Korean	10.0	7.752922	1.260174	5.602794	7.125573	7.798156	8.761087	9.486981
	Spanish	10.0	7.435105	1.939005	5.077368	6.107491	6.931256	9.082138	10.536023
5	Arabic	10.0	7.138747	1.646513	4.248529	5.888579	7.525306	8.581431	8.870455
	Korean	10.0	6.803422	1.714505	4.274374	5.384871	7.081285	7.685704	9.499716
	Spanish	10.0	7.891529	1.291962	5.048252	7.536954	8.304324	8.613962	9.712147

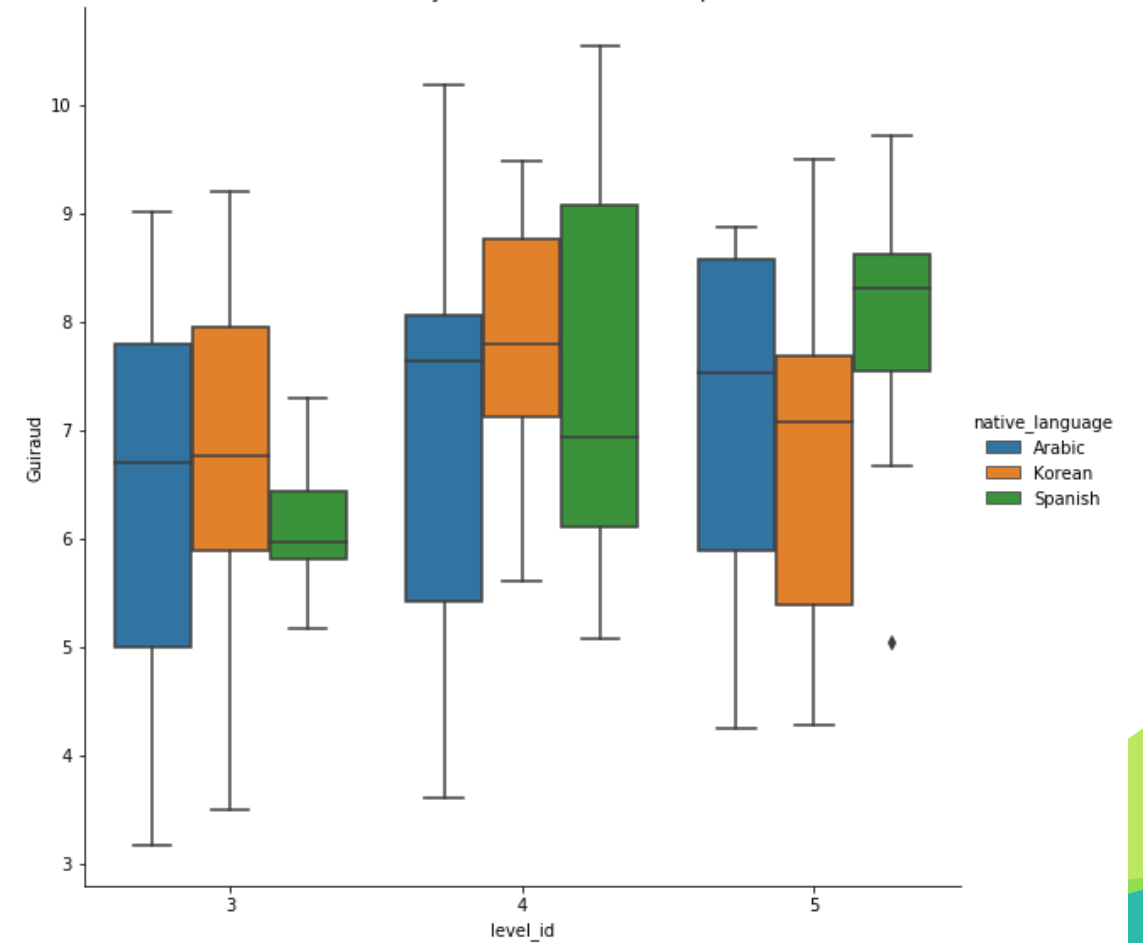# Samples

BALC

Guiraud's R

PELIC

Guiraud's R by Level (BALC samples)
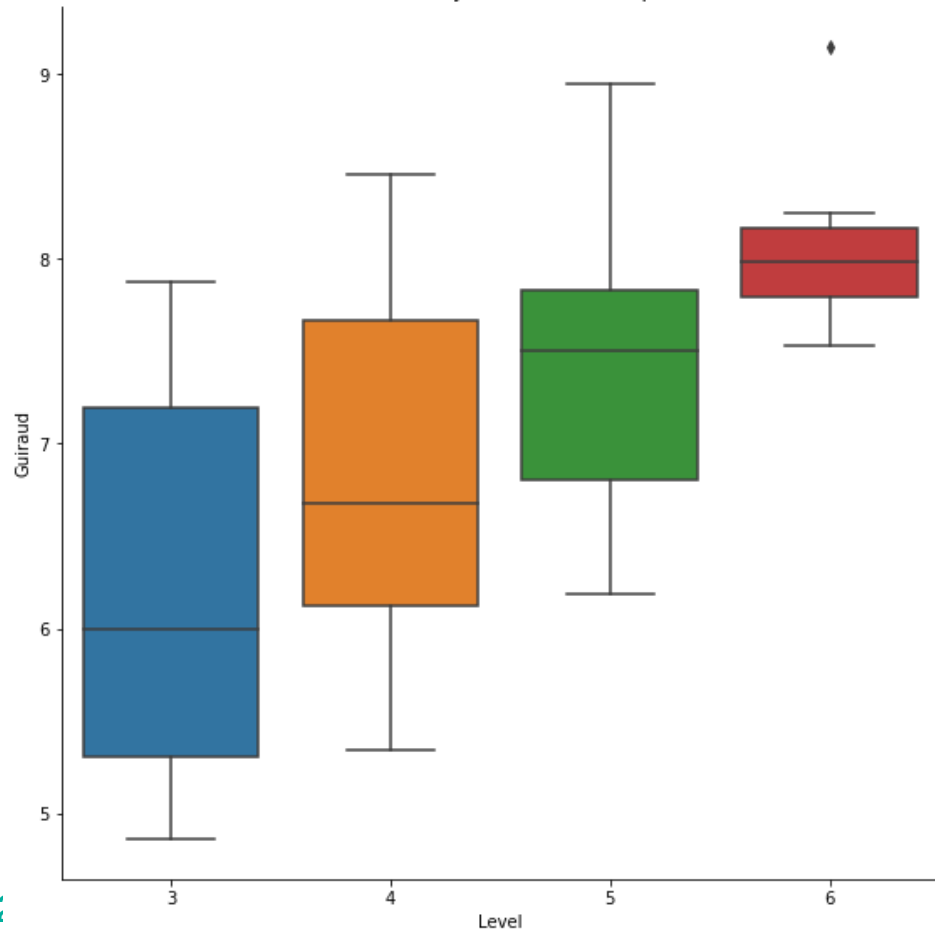
Guiraud's R by Level and L1 (PELIC samples)

Analysis (PELIC)

PELIC, Arabic 3 ('book'):

Topic: personal success
Making a dessert

I think I have a personal success in making a lot of traditional desserts from several countries. There were some reasons for helping my superior success. One of the reasons, I gained this talent from my mother. She was a wonderful cook especially when she made cakes. In fact, my mother taught me how to prepare German desserts. For example, I learned the crepes, many kinds of the nut cakes, fruit cakes, and the cake's topping, too. Another reason for my success, I read many of national and international dessert books for several times ago. I practiced the desirable recipes as soon as I read about them. Finally, I've disciplined myself to watch the food channel on T.V for half an hour a day. I feel that some of my favorite food channels would help me to improve my talent in the last two years.

Analysis (PELIC)

PELIC, Korean 3 ('book'):

I have little furniture in my room.
As you enter the door through the entrance, desk is in front of you. The desk is on the front, between book chests. I have two book chests. One book chest is on the right wall. The other one is on the front wall. My bed is on the left corner, beside book chest. Clothes chest is on the left, in front of bed. My room center has the heart shape rug.

Analysis (PELIC)

PELIC, Spanish 3 ('book'):

There are three types of magazines, they are available from all good bookshops or online. I can to get the best preference or the information about if it is interesting for me.

The first type of magazine is a sport magazine. For example, I can find magazine providing the latest information on building and operating an athletic facility, athletic programs, action photos, athlete interviews, scores and statistic on an professional and college sports. The second type of magazine is an automobile whether if I have a passion for cars, I can find a long list of magazines especially tailored for the information needs of car such as car design, car market, or car lifestyle. The last type is a health magazine these cover a variety of topics including physical fitness, nutrition, beauty, weight training and body building.
In conclusion, I can find many types of magazines the most important, I can choose or find my preference and what information is interesting for me.

Analysis (BALC)

BALC 3:

Last summer holiday iam going with my family to perfect weekend. we going to the India. We going to wooke in the booke and go to shopping. And we will go to see Tag Mhale. It was so wonderful. cause it have a nice people and a nice shope. After I finsed in India iam going to Hangare with my family. we spend in Loxx. And iam plaing football and going to swam.

BALC 4:

I have just had the perfect holidy. In my holiday I went to dubai, Abu dabi, Sharjah Al-Ain and Fujirah I went there places with my family and my friends. I went the cinama and watched nice movies with my friends, I went malls I buys clothes and play some time with sisters and brothers for example went the sea with my family I just had a wonderfor time I swam in the sea and play with my father and sisters football. I'm very happy becouse I have nice holiday. In a some time I went net coffe set in a nel and drank some coffe or juce. On day in aholiday I went the my friends house she wented your friend reading some Bookes, magazims, and stores the mother my friend set and to took for a what's to do in holiday she gived nice idea. open the eyes and close the eyes and the butiful holidy finishd The holiday is wonderful becouse I went many places with my family and my friends.

Analysis (BALC)

BALC 5:

In the summer holiday my family dicides to spend our holiday in Japan , so I prepared my self to spend 2 week there. When I reached Japan ,I feel so intresting when I saw Japanese child wearing traditional dress(keymono) and playing with water. The weather there was so nice . We stay in small house made of wood and there was a cute family from Japan living with us. They was so kind with us. They took us to have a nice Jeirney in the city. When I reached the city I been shocked when I saw people every where. there holding magazine or newspaper or story book, every where on the bus, in the street on the park every where, every where they Just read and work all the time. They really had diffrent life. When our Jeirney finished in the city .they took us to Japanese hotel where I eat a deliscious food. Really I spend intresting time in Japan I learned some Japanese word and I learned about there culture and tradition, and I learned how to manage my time. Really it was a nice and perfect holiday I have ever spend.

BALC 6:

There have been many films reteased lately, Some were good and others were awful.. but there's always that little group of films that shine through the rest, films that are so good that you don't mind watching them a second time. The film that I enjoyed the most was based on and named after the classic novel "Pride and Prejudice" It stars Keira Knightley who's one of my favourite actresses. The story is about how important first impressions are and how they effect the way aperson is seen. This happened when Elizabeth bennet overhears Mr.Darsey , Arich gentleman complaing to his friend about the party they were attending and how he thought that Elizabeth was not pretty and called her plain. After that Elizabeth thinks of Mr. Darsey as an arrogant man and doesn't want anything to do with him, She also refuses his marrige proposal because of her pride, And it takes many incedents and some time before she realises that she had fallen in love with him and the film ends with her accepting his proposal.. I enjoyed this film because of the great story, the wonderful scenery and because it can be seen in the real world , last but not least I learnt that we shouldn't judde a book by it's cover .

Thank you!

Questions?

References

◆ Butler, Y. G. (2002). Second language learners' theories on the use of English articles. *Studies in Second Language Acquisition, 24*, 451-480.

◆ Ionin, T., Zubizarreta, M. L. & Maldonado, S. B.. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua, 118*, 554-576.

◆ Master, P. (1997). The English article system: acquisition, function, and pedagogy. *System, 25*, 215-232.

◆ Stehle, M. E. (2009). *The Interlanguage Development Of Articles In English As A Second Language: A Longitudinal Study*. [University of Pittsburgh ETD].

◆ Zhao, H. &MacWhinney, B. (2018). The Instructed Learning of Form–Function Mappings in the English Article System. *The Modern Language Journal, 102*(1), 99-119.

◆ Zobl, H. (1982). A direction for contrastive analysis: The comparative study of developmental sequences. *TESOL Quarterly, 16*, 169-183.