Experiments on the pages of Cognitive Linguistics from 2012 to 2022

Jane Klavan, Mariann Proos, Piia Taremaa, Ann Veismann

The call for empiricism within cognitive linguistics was launched more than 25 years ago
(e.g. Sandra and Rice 1995; Cuyckens et al. 1997)

10 years ago Laura Janda wrote that both the field of cognitive linguistics as a whole and the journal Cognitive Linguistics have taken a quantitative turn
(Janda 2013)

Janda (2013) surveyed all of the articles published in the journal from its first volume in 1990 through to the volume of 2012 and observed an exponential growth in studies that use statistical analysis of corpus data and experimental findings.

We want to follow up on the original survey conducted by Janda (2013) by looking at the articles published in the journal *Cognitive Linguistics* from 2012 to 2022.

In our systematic review, we will focus mainly on experimental methods.

- We expect the number of papers using experimental methods to have risen over the years
  - *inter alia*, we expect the rise in more complex research designs and more advanced statistical modelling techniques
- cf. some of the methodological discussions that have taken place in the field of linguistics in general (e.g. Dąbrowska 2010, 2016; Edelman and Christiansen 2003; Gibson and Fedorenko 2010, 2013; Grieve 2021; Sprouse and Almeida 2013).

Question 1

What is the relative proportion of using experiments in comparison with other methods, e.g. corpus analysis, for conducting cognitive linguistic research?

Question 2

Are some experimental designs used more often and hence deemed more suitable to answer the types of research questions cognitive linguists are interested in?

Data

Year	Articles	Year	Articles
2012	20	2018	21
2013	20	2019	26
2014	22	2020	22
2015	19	2021	22
2016	24	2022	23
2017	22	Total	241

- DOI: DOI of the article
- Year: the year of publication
- Issue: the issue of publication
- Method: corpus study, experimental study, combination qualitative study
- Experiment_yes/no: Does the study report the findings of an experiment? - Yes/No
- Type of experiment: Type of experiment as described in the article
- N of participants: Number of participants
- N of test items: Number of test items
- Keywords: List of keywords copied from the article
- Language(s): Language(s) that have been studied

- Janda (2013) surveyed all of the articles published in *Cognitive Linguistics* from its inaugural volume in 1990 through to 2012.
- A total of 331 articles were surveyed (excluding review articles, book reviews, overviews, commentaries, replies, squibs, CLiPs, introductions to special issues)
- She defined a "quantitative article" as an article in which "a researcher reports numbers for some kind of authentic language data" (Janda 2013: 4)
- 141 quantitative articles during 1990-2012 (42%)

Figure 1. *Percent quantitative articles in* Cognitive Linguistics 1990–2012
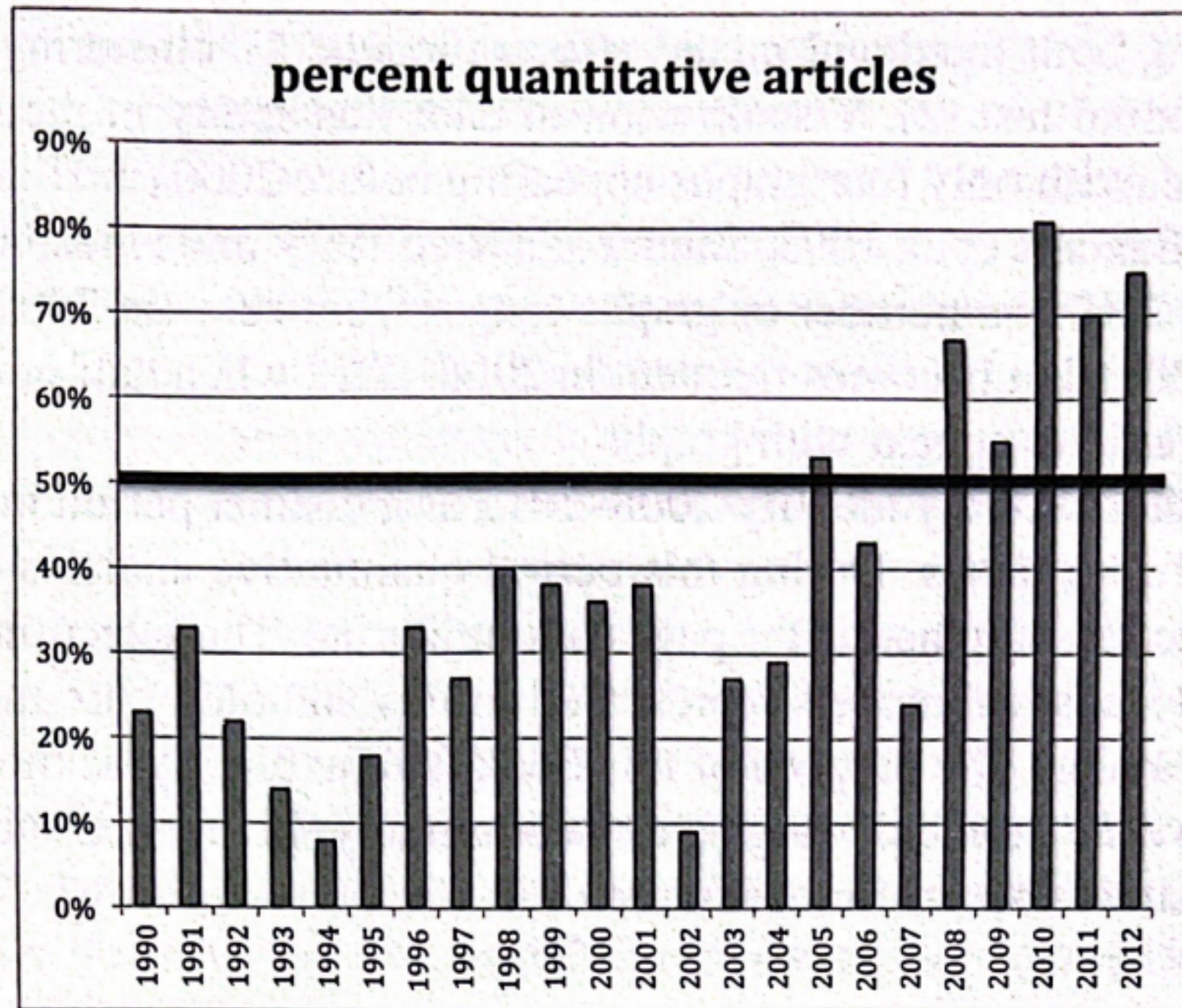
Janda (2013: 4-5) divides the history of *Cognitive Linguistics* into two eras:

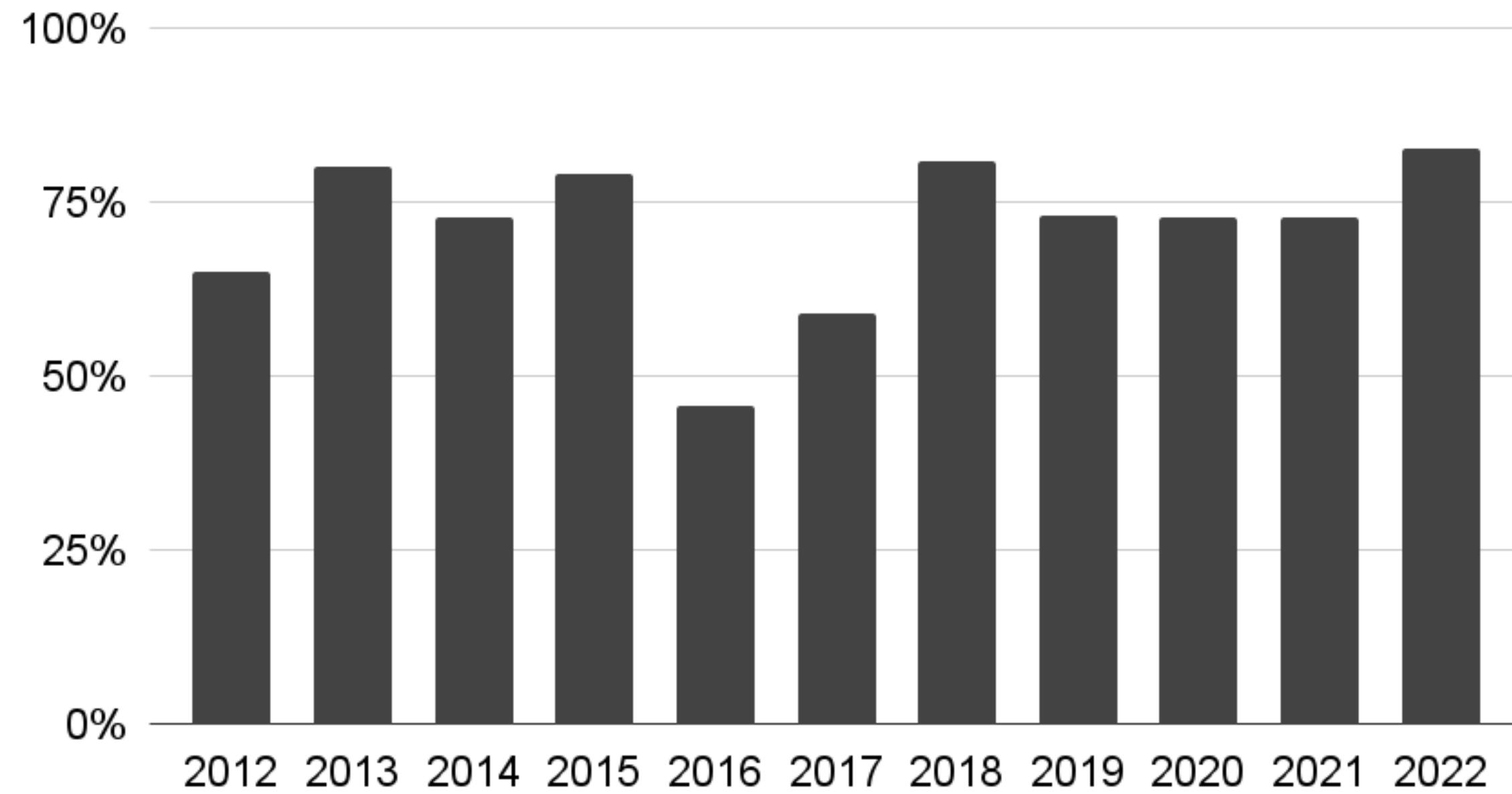1990-2007 - most articles were not quantitative

2008-2012 most articles were quantitative

- "The majority of quantitative articles in our journal report corpus data (34%) or experimental data (48%) or a combination of the two (6%), and acquisition data (which can involve both corpus and experimental data) is also steadily represented (12%)." (Janda 2013: 5)
- "We can thus securely identify 2008-2012 as a distinct period in the history of *Cognitive Linguistics*. During this period quantitative analysis emerges as common practice, dominating the pages of our journal." (Janda 2013: 6)
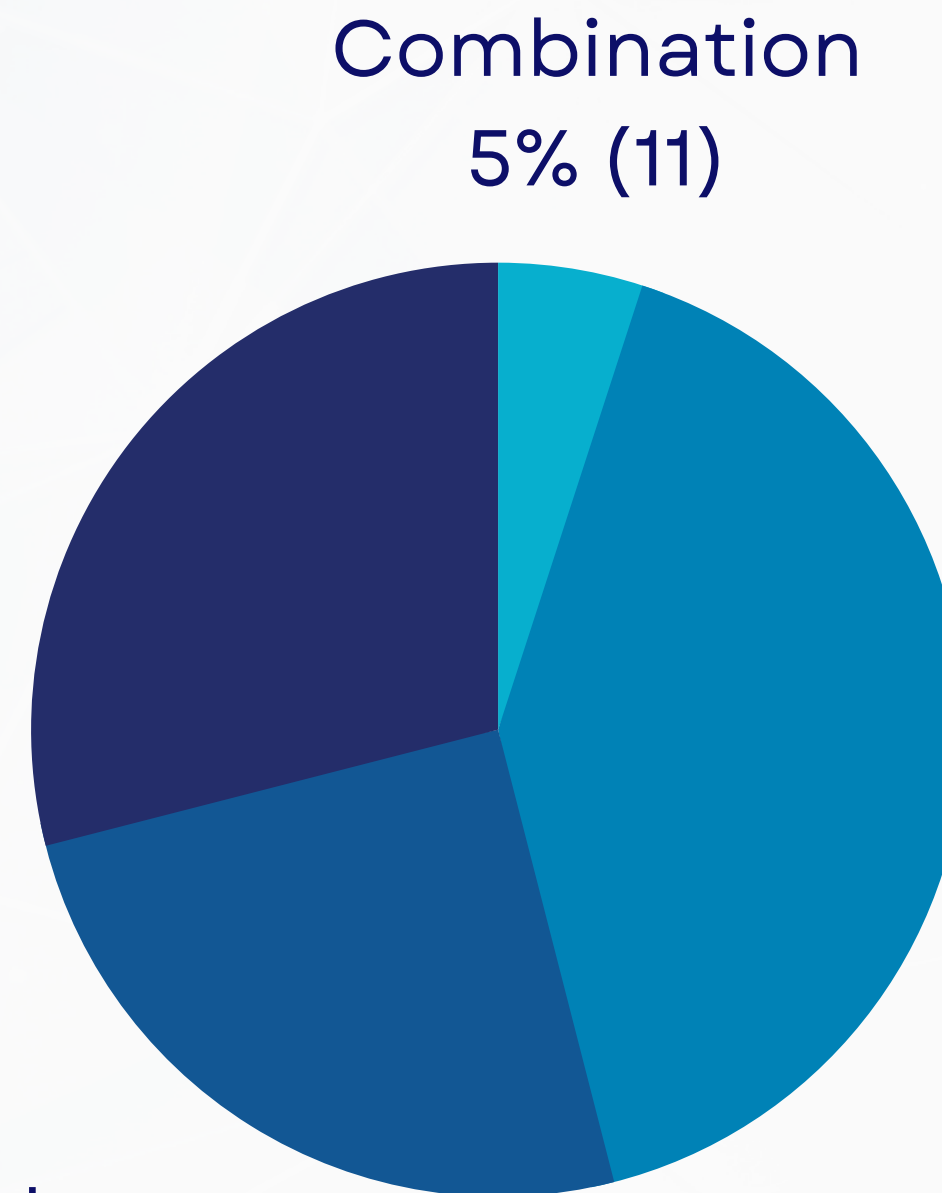
percent quantitative articles

Question 1

What is the relative proportion of using experiments in comparison with other methods, e.g. corpus analysis, for conducting cognitive linguistic research?

% of articles reporting experimental data
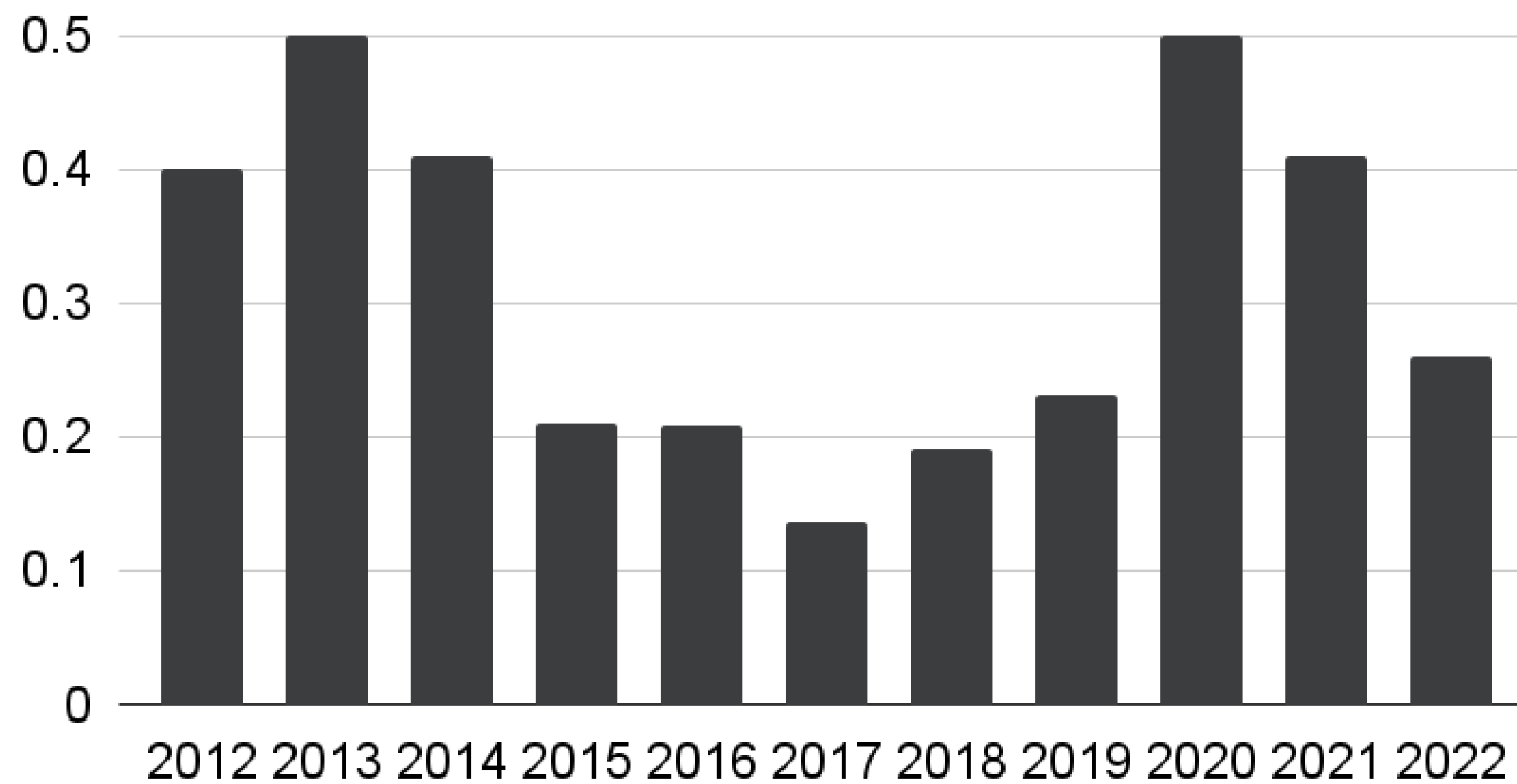
Question 2

Are some experimental designs used more often and hence deemed more suitable to answer the types of research questions cognitive linguists are interested in?

Table 1.   *Kinds of linguistic data (sorted according to naturalness of production/collection)*

#	Data source
1	corpora with written texts (e.g., newspapers, webblogs)
2	example collections
3	corpora of recorded spoken language in societies/communities where note-taking/recording etc is not particularly spectacular/invasive
4	corpora with recorded spoken language from fieldwork in societies/communities where note-taking/recording etc is spectacular/invasive
5	data from interviews (e.g., sociolinguistic interviews)
6	experimentation requiring subjects to do something with language they usually do anyway, e.g., – tsentence production as in answering questions in studies on priming – tpicture description in studies on information structure
7	elicited data from fieldwork (e.g., responses to "how do you say X in your language?")

8 experimentation requiring subjects to do something with language they usually do not do, on units they usually interact with, e.g.,
   – sentence sorting
   – measurements of reaction times in lexical decision tasks
   – word associations

9 experimentation requiring subjects to do something with language they usually do not do

   on units they usually interact with, involving typical linguistic output, e.g.,
   – measurements of event-related potentials evoked by viewing pictures
   – eye-movements during reading idioms
   – acceptability/grammaticality judgments
   on units they usually do not interact with, involving the production of linguistic output, e.g.,
   – phoneme monitoring
   – gating
   – ultrasound tongue-position videos

Additional question

What languages have been studied on the pages of *Cognitive Linguistics* from 2012 - 2022?

Our results

Languages covered in *Cognitive Linguistics* (2012-2022) over 66 different languages

Language	Number of experimental studies	Total number of studies		Language	Number of experimental studies	Total number of studies
English	31	100		Äiwoo	1	2
Chinese	4	8		Arabic	1	1
Dutch	4	15		ASL	1	8
ESL	4	5		Bashkir	1	1
Japanese	4	7		Danish	1	1
Mandarin	4	9		French	1	8
German	3	9		Hungarian	1	2
Spanish	3	13		Italian	1	1
Finnish	2	3		ItalianSignLanguage	1	1
Polish	2	5		Norwegian	1	1
Russian	2	6		Ryukyuan	1	1
				Secoya	1	1
				Serbian	1	1
				Swedish	1	1
				Thai	1	1

Experiments on the pages of Cognitive Linguistics from 2012 to 2022

01 Over 70% of articles are quantitative = 41% corpus data, 25% experimental data, 5% combination of both (cf. Janda 2013: ~20% experimental data)

02 Most frequent experimental designs include: elicitation task, rating task, forced choice task, survey, eye-tracking, self-paced reading experiment

03 Over 66 different languages have been studied; the most frequent ones are Indo-European languages (English, Dutch, Spanish, German, ASL)

04 Sometimes, it is surprisingly difficult to understand what exactly the researchers have studied and how - methodology, participants, items, etc.

There are naturally those within the field who claim that there is too much "number-crunching" and too little introspective (qualitative) research being done (Langacker 2016), and those who claim that the field is still very much dependent on introspective data and not enough empirical research is being done (Dąbrowska 2016).

Larsson, T., Egbert, J., & Biber, D. (2022). On the status of statistical reporting versus linguistic description in corpus linguistics: A ten-year perspective.

"... research articles published in four major corpus linguistics journals in 2009 and 2019. The results display a marked change: in 2009, a clear majority of the articles exhibit a preference for linguistic description over statistical reporting; in 2019, the exact opposite is true."

"Whilst the increased statistical focus may reflect increased methodological sophistication, our results show that it has come at a cost: a diminished focus on linguistic description, evident, for example, through fewer text excerpts and linguistic examples, which appears to be symptomatic of increasing distance from the language that is the object of study."

"A linguist can bring speakers of different languages or dialects into the lab to compare some aspect of language production or perception, but such analyses will necessarily be observational, as the linguist cannot directly control the social background of speakers."

"For this reason, linguistics has traditionally been an observational field of study. Observation is the basis of many modern fields of linguistics, [...] allowing for important theories and models of language to be developed and assessed, generally without any reliance on experimentation."

- Justification for the choice of the journal Cognitive Linguistics (cf. Janda 2013) - it gives us the most consistent perspective available on the use of experiments in the field.
- Cognitive linguists who are using experimental work in their research may choose other venues for publishing their research.
- Our aim was to give an overview of the situation as it is portrayed on the pages of the "official" journal.

In our future work, we want to extend the survey to include other prominent venues for cognitive linguists.

- Is there a rise in more complex research designs and more advanced statistical modelling techniques?
- What are the implications and conclusions drawn by researchers applying different methods - every method counts but for what?

Fostering the discussion on the importance of methodological decisions and what these decisions entail in terms of interpreting the data and building cognitive linguistic theories.

References

- Cuyckens, Hubert, Dominiek Sandra and Sally Rice. 1997. Towards an empirical lexical semantics. In Birgit Smeija and Meike Tasch (eds). Human Contact Through Language and Linguistics, 35–54. Bern: Peter Lang.
- Dąbrowska, Ewa. 2010. Naive v. expert intuitions: an empirical study of acceptability judgements. The Linguistic Review, 27: 1, 1–23.
- Dąbrowska, Ewa. 2016. Cognitive Linguistics' seven deadly sins. Cognitive Linguistics, 27: 4, 479–491.
- Edelman, Shimon, and Morten H. Christiansen. 2003. How seriously should we take minimalist syntax? A comment on Lasnik. Trends in Cognitive Science, 7: 2, 60–61.
- Gibson, Edward, and Evelina Fedorenko. 2010. Weak quantitative standards in linguistics research. Trends in Cognitive Sciences, 14: 6, 233–234.
- Gibson, Edward, and Evelina Fedorenko. 2013. The need for quantitative methods in syntax and semantics research. Language and Cognitive Processes, 28: 1, 88–124.
- Grieve, Jack. 2021. Observation, experimentation, and replication in linguistics. Linguistics, 59: 5, 1343–1356.
- Janda, Laura A. (ed.). 2013. Cognitive linguistics: The quantitative turn. Berlin: De Gruyter Mouton.
- Langacker, Ronald W. Working toward a synthesis. Cognitive Linguistics, 27: 4, 465–477.
- Sandra, Dominiek and Sally Rice 1995. Network analyses of prepositional meaning: mirroring whose mind – the linguist's or the language user's? Cognitive Linguistics, 6: 1, 89–130.
- Sprouse, Jon, and Diogo Almeida. 2013. The empirical status of data in syntax: A reply to Gibson and Fedorenko. Language and Cognitive Processes, 28: 3, 222–228.